

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:23:06 ; Search time 11.0618 Seconds
(without alignments)
476.116 Million cell updates/sec

Title: US-09-508-487-19
Perfect score: 902
Sequence: 1 MNKLLIFVLATFCVFSFAQ.....DVAMGSSALGFELSFKKS 179

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------------------------|
| 1 | 80 | 8.9 | 243 | 4 | US-09-134-001C-2961 Sequence 2961, Ap |
| 2 | 77.5 | 8.6 | 392 | 4 | US-09-071-035-260 Sequence 260, App |
| 3 | 77.5 | 8.6 | 513 | 4 | US-09-291-922-20 Sequence 20, Appl |
| 4 | 77.5 | 8.6 | 1638 | 4 | US-09-071-035-258 Sequence 258, App |
| 5 | 77.5 | 8.6 | 1638 | 4 | US-09-071-035-262 Sequence 262, App |
| 6 | 77.5 | 8.6 | 1638 | 4 | US-09-071-035-266 Sequence 266, App |
| 7 | 74.5 | 8.3 | 222 | 4 | US-09-134-001C-3521 Sequence 3521, Ap |
| 8 | 72 | 8.0 | 332 | 1 | US-08-118-270-53 Sequence 53, Appl |
| 9 | 72 | 8.0 | 332 | 5 | PCT-US93-08528-53 Sequence 4385, Ap |
| 10 | 71.5 | 7.9 | 300 | 4 | US-09-134-001C-4385 Sequence 3009, Ap |
| 11 | 71.5 | 7.9 | 303 | 4 | US-09-134-001C-3009 Sequence 5331, Ap |
| 12 | 71.5 | 7.9 | 655 | 4 | US-09-134-001C-5531 Sequence 22, Appl |
| 13 | 71.5 | 7.9 | 858 | 4 | US-09-255-829-22 Sequence 29, Appl |
| 14 | 71.5 | 7.9 | 858 | 4 | US-09-255-829-29 Sequence 20, Appl |
| 15 | 71.5 | 7.9 | 1169 | 4 | US-09-255-829-20 Sequence 7, Appl |
| 16 | 71 | 7.9 | 266 | 2 | US-08-225-480-7 Sequence 4, Appl |
| 17 | 71 | 7.9 | 266 | 2 | US-09-097-759-4 Sequence 7, Appl |
| 18 | 71 | 7.9 | 266 | 4 | US-09-118-445-7 Sequence 3, Appl |
| 19 | 71 | 7.9 | 299 | 2 | US-09-097-759-3 Sequence 2, Appl |
| 20 | 71 | 7.9 | 306 | 2 | US-09-097-759-2 Sequence 6, Appl |
| 21 | 70.5 | 7.8 | 402 | 1 | US-08-314-309A-6 Sequence 4, Appl |
| 22 | 70.5 | 7.8 | 441 | 1 | US-08-314-309A-4 Sequence 4, Appl |
| 23 | 70.5 | 7.8 | 696 | 4 | US-07-757-342D-4 Sequence 4004, Ap |
| 24 | 70 | 7.8 | 349 | 4 | US-09-134-001C-4004 Sequence 16, Appl |
| 25 | 70 | 7.8 | 612 | 3 | US-09-295-186-16 Sequence 3487, Ap |
| 26 | 69.5 | 7.7 | 477 | 4 | US-09-134-001C-3487 Sequence 5, Appl |
| 27 | 69 | 7.6 | 269 | 1 | US-08-447-554-5 |

| | | | | | |
|----|------|-----|------|---|---------------------------------------|
| 28 | 69 | 7.6 | 269 | 1 | US-08-468-763-17 Sequence 17, Appl |
| 29 | 69 | 7.6 | 269 | 1 | US-08-448-160-5 Sequence 5, Appl |
| 30 | 69 | 7.6 | 269 | 1 | US-08-393-996A-17 Sequence 17, Appl |
| 31 | 69 | 7.6 | 611 | 4 | US-07-757-342D-8 Sequence 8, Appl |
| 32 | 69 | 7.6 | 636 | 4 | US-07-757-342D-7 Sequence 7, Appl |
| 33 | 69 | 7.6 | 674 | 4 | US-07-757-342D-10 Sequence 10, Appl |
| 34 | 69 | 7.6 | 699 | 4 | US-07-757-342D-2 Sequence 2, Appl |
| 35 | 69 | 7.6 | 1861 | 2 | US-08-790-912-4 Sequence 4, Appl |
| 36 | 68.5 | 7.5 | 494 | 4 | US-09-134-001C-4475 Sequence 4475, Ap |
| 37 | 68 | 7.5 | 462 | 2 | US-08-898-976-2 Sequence 2, Appl |
| 38 | 68 | 7.5 | 462 | 2 | US-08-898-976-4 Sequence 4, Appl |
| 39 | 68 | 7.5 | 831 | 2 | US-08-677-734A-11 Sequence 11, Appl |
| 40 | 68 | 7.5 | 831 | 4 | US-09-097-053-11 Sequence 11, Appl |
| 41 | 67.5 | 7.5 | 682 | 2 | US-08-436-900A-4 Sequence 4, Appl |
| 42 | 67.5 | 7.5 | 834 | 2 | US-08-677-734A-9 Sequence 9, Appl |
| 43 | 67.5 | 7.5 | 834 | 2 | US-08-677-734A-10 Sequence 10, Appl |
| 44 | 67.5 | 7.5 | 834 | 4 | US-09-097-053-9 Sequence 9, Appl |
| 45 | 67.5 | 7.5 | 834 | 4 | US-09-097-053-10 Sequence 10, Appl |

ALIGNMENTS

RESULT 1
US-09-134-001C-2961
; Sequence 2961, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2961
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2961

Query Match 8.9%; Score 80; DB 4; Length 243;
Best Local Similarity 30.3%; Pred. No. 0.18;
Matches 36; Conservative 16; Mismatches 53; Indels 14; Gaps 5;
QY 60 GIGSFAOGDILGGSLILGFDVAGLILAGAYLDIKALDITKKAFOWTGKGVMLAGV 119
Db 16 GIMSFSFLITAGVAAFNPCGIAL--PSTISY-LIGGETKDHISFRIATFKGLGLGA 72
QY 120 VTMAVTRLTETILPFTFANSYNRKLNSLVALGFGFSDVAMGQSSA-LGFELSFKK 177
Db 73 MT-----TGELTIFVLG----LLIGGLSALTGFIFFILSLVMGILIALLLGLMFGK 121

RESULT 2
US-09-071-035-260
; Sequence 260, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 260:
SEQUENCE CHARACTERISTICS:
LENGTH: 392 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-260

Query Match 8.6%; Score 77.5; DB 4; Length 392;
Best Local Similarity 27.7%; Pred. No. 0.68;
Matches 43; Conservative 22; Mismatches 59; Indels 31; Gaps 8;
QY 21 ANDSKN-----GAFGMSAGEKLLVYETSKODPIVPPFLNLFLGFGIGSFAQGDILG-- 71
DB 106 ANDLNTVITQITSGALGNLTGVDIDLTEVNRQLELVNNTIENL-----GAASFTAPETLAAD 161
QY 72 GSLILGFDVAGIGLILAGAYLDIKALDGIITKKAFTWTKGVMLAGVVTMAVTRLTEII 131
DB 162 GSYISAPISDGLGLVLAQNVSNI--LQDL--NAAVQALEAKGTSIPS--NLVAAAANAAL 215
QY 132 LPFTFANSYNRKLNLSNALGSGFEPSPDVANGQS 166
DB 216 LP-----VKGTNVNAVSGALPL--LAVGGS 238

RESULT 3
US-09-291-922-20
Sequence 20, Application US/09291922
Patent No. 6383776
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Hitz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Tingey, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: BB-1163
CURRENT APPLICATION NUMBER: US/09/291,922
EARLIER FILING DATE: 1999-04-14
EARLIER APPLICATION NUMBER: 60/083,044
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Microsoft Office 97
SEQ ID NO 20
LENGTH: 513
TYPE: PRP
ORGANISM: Zea mays
US-09-291-922-20
Query Match 8.6%; Score 77.5; DB 4; Length 513;
Best Local Similarity 26.0%; Pred. No. 0.98;
Matches 32; Conservative 18; Mismatches 30; Indels 43; Gaps 5;

QY 3 KLLIFVLATFCVSSFAQANDSKNGAFGMSAGEKLLVYETSKODPIVPPFLNLFLGFGIG 62

DB 61 EILIGILSYLSFGSPAGARTSDR-----IGRRLTVV----- 92
QY 63 SFAQGDILGSLILGFDVAGIGLILAGAYLDIKALDGIITKKAFTWTKGVMLAGVVTM 122
DB 93 -FAAVIFFVGSLLMGF-AVNYGMLMAGREV---AGGV-----GYGGMIAPIVYTA 137
QY 123 AVT 125
DB 138 EIS 140
RESULT 4
US-09-071-035-258
Sequence 258, Application US/09071035
Patent No. 6448043
GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 258:
SEQUENCE CHARACTERISTICS:
LENGTH: 1638 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-258

Query Match 8.6%; Score 77.5; DB 4; Length 1638;
Best Local Similarity 27.7%; Pred. No. 5;
Matches 43; Conservative 22; Mismatches 59; Indels 31; Gaps 8;
QY 21 ANDSKN-----GAFGMSAGEKLLVYETSKODPIVPPFLNLFLGFGIGSFAQGDILG-- 71
DB 151 ANDLNTVITQITSGALGNLTGVDIDLTEVNRQLELVNNTIENL-----GAASFTAPETLAAD 206
QY 72 GSLILGFDVAGIGLILAGAYLDIKALDGIITKKAFTWTKGVMLAGVVTMAVTRLTEII 131
DB 207 GSYISAPISDGLGLVLAQNVSNI--LQDL--NAAVQALEAKGTSIPS--NLVAAAANAAL 260
QY 132 LPFTFANSYNRKLNLSNALGSGFEPSPDVANGQS 166
DB 261 LP-----VKGTNVNAVSGALPL--LAVGGS 283

RESULT 5


```
Db 59 FGKLSGVITIANIOTVPVIAMLAILMSGLGM-----NTVFTVFLVALLPIIK 110
QY 89 GAYLDKALDKITKKAQFQWTKGYVLAGVVTMAVTR---LREIILPFTFANSYNRKLK 145
Db 111 NNTYGINEVDPNPKDA-----GKG-----MGTRNQVLTWIELPL----- 145
QY 146 NSLNVALGGFPEFSDVAMG 164
Db 146 -SLSVIIGGIRIALVVAIG 163

RESULT 8
US-08-118-270-53
; Sequence 53, Application US/08118270
; Patent No. 5508384
; GENERAL INFORMATION:
; APPLICANT: Murphy, Randall B.
; APPLICANT: Schuster, David I.
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/118,270
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY-2A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-118-270-53

Query Match 8.0%; Score 72; DB 1; Length 332;
Best Local Similarity 28.7%; Pred. No. 2.4;
Matches 41; Conservative 19; Mismatches 41; Indels 42; Gaps 8;

QY 37 LLVYETSKODPIVP-FLNLFLGFGIGSFAQGDILGSLILGFDVAVGIGLILAGAYLDIK 95
Db 23 LFLVLLTSRYKLTVPREIMNL-----SFA-----DFCMLYLL-----IA 56
QY 96 ALDGITK-----KAAFWTKGKGVMLAGVVT-----MAVTRLEIIL-----PPTFANSYNR 142
Db 57 SVDSQTKGYYNHAIQWQSGCSTAGFTVLASELSVTLTVITLERWHTITYAIHQ 116
QY 143 K--LKNSLNVALGGFPEFSDVAM 163
Db 117 KLRLRAILMLGGWLFSSLIAM 139

US-08-118-270-53

Query Match 8.0%; Score 72; DB 1; Length 332;
Best Local Similarity 28.7%; Pred. No. 2.4;
Matches 41; Conservative 19; Mismatches 41; Indels 42; Gaps 8;

QY 37 LLVYETSKODPIVP-FLNLFLGFGIGSFAQGDILGSLILGFDVAVGIGLILAGAYLDIK 95
Db 23 LFLVLLTSRYKLTVPREIMNL-----SFA-----DFCMLYLL-----IA 56
QY 96 ALDGITK-----KAAFWTKGKGVMLAGVVT-----MAVTRLEIIL-----PPTFANSYNR 142
Db 57 SVDSQTKGYYNHAIQWQSGCSTAGFTVLASELSVTLTVITLERWHTITYAIHQ 116
QY 143 K--LKNSLNVALGGFPEFSDVAM 163
Db 117 KLRLRAILMLGGWLFSSLIAM 139
```

```
RESULT 9
PCT-US93-08528-53
; Sequence 53, Application PC/TUS9308528
; GENERAL INFORMATION:
; APPLICANT: New York University
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08528
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY-2 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US93-08528-53

Query Match 8.0%; Score 72; DB 5; Length 332;
Best Local Similarity 28.7%; Pred. No. 2.4;
Matches 41; Conservative 19; Mismatches 41; Indels 42; Gaps 8;

QY 37 LLVYETSKODPIVP-FLNLFLGFGIGSFAQGDILGSLILGFDVAVGIGLILAGAYLDIK 95
Db 23 LFLVLLTSRYKLTVPREIMNL-----SFA-----DFCMLYLL-----IA 56
QY 96 ALDGITK-----KAAFWTKGKGVMLAGVVT-----MAVTRLEIIL-----PPTFANSYNR 142
Db 57 SVDSQTKGYYNHAIQWQSGCSTAGFTVLASELSVTLTVITLERWHTITYAIHQ 116
QY 143 K--LKNSLNVALGGFPEFSDVAM 163
Db 117 KLRLRAILMLGGWLFSSLIAM 139

RESULT 10
US-09-134-001C-4385
; Sequence 4385, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
```


;; PRIOR FILING DATE: 1997-11-08
;; PRIOR APPLICATION NUMBER: US 60/055,779
;; PRIOR FILING DATE: 1997-08-14
;; NUMBER OF SEQ ID NOS: 5674
;; SEQ ID NO 4385
;; LENGTH: 300
;; TYPE: PRT
;; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4385

Query Match 7.9%; Score 71.5; DB 4; Length 300;
Best Local Similarity 18.8%; Pred. No. 2.4;
Matches 51; Conservative 30; Mismatches 67; Indels 123; Gaps 11;
QY 1 MNKLIFVLATFCVFFSFAQANDSKNG-AFGMSAGEKLLVYETS----- 43
DB 2 MRKVLIFAIAGF-----LAQLVDGSLGMSGFACSSILLIYGIAPVVSATVHFSEIATT 56
QY 44 -----KODPI-----VPLLNLFLGFGIGSFAQGD----- 68
DB 57 AASGTSHWRDENVHKPTMLKLAIPGSIASAFAGAGVLFHGDYIKPFIALFLLSMGFYIL 116
QY 69 -----ILGSLILGF-DAVGIG-----LILAGAVLDIKAL 97
DB 117 YQFLFKRAHEHHHVGNLSSFKVPIQGFVAGFLDAIGGGWGPVNTPLLSKKKIOPRYA 176
QY 98 DG-----ITKKAFAQWTWKGK-----MLAGVVTMAVTRLEIILPFTFAN 138
DB 177 IGTVSASEFFVTSSAALSFIIFLGVTQINWFAVIALSLGGMVAAPISAYLVKVLPI----- 232
QY 139 SYNKRKLNSLNLVALGFGFSPFDVAMGQSSAL 169
DB 233 -----NILAICVGGI-----IIFTNSNAL 251

RESULT 11
US-09-134-001C-3009
; Sequence 3009, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3009
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3009

Query Match 7.9%; Score 71.5; DB 4; Length 303;
Best Local Similarity 21.0%; Pred. No. 2.4;
Matches 41; Conservative 31; Mismatches 66; Indels 57; Gaps 9;
QY 1 MNKLIF-----VLATFCVFS-----SFAQANDSKNGAFGMSA 33
DB 108 MNPLAVFMLVLLGWKVAIVYFVIAIFSIPTGLVFSKMNLAEYKGVNWKDGFDFRANKE 167
QY 34 GEKLLVYETSKQDP---IVPFLNLFLGFGIGSFAQGD-----LGSLILGFDAV 81
DB 168 GSR---FKQALNDAWAFYPLMPLVLFYGVFGVFIYGFVPEFTFTKYASGDGVISVFIGS 224
QY 82 GIGLILAGVLDIKALDGIYTKKAQFQWTKGVMAGVVTMAV-----TRITEIILPFTFA 137
DB 225 VIGIPM---YIRPETMLPIAEALV-----SKGMSLGTWVALLIGGASGASIPVVL---LT 273

QY 138 NSYNRKLKNSLVAL 152
DB 274 KLFKKKFWSEFVIAI 288
RESULT 12
US-09-134-001C-5531
; Sequence 5531, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5531
; LENGTH: 655
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5531

Query Match 7.9%; Score 71.5; DB 4; Length 655;
Best Local Similarity 24.4%; Pred. No. 7;
Matches 42; Conservative 29; Mismatches 74; Indels 27; Gaps 7;
QY 15 FSSFAQA--NDSKNGAFGMSAGEKLLVYETSKQDPVFPFLNLFLGFGIG---SFAQGD 69
DB 349 YNAFQEQLWNIGKNSAFAL-----IIPILAG-FIARSIADKPGFAG-L 390
QY 70 LGGSLILGFDAVGIGLILAG---AYLDIKALDGIYTKKAQFQWTKGVMAGVVTMAVTR 126
DB 391 VGGMLAISGGSGFGIGIAGFLAGYL-TQGIKYITRKLPOAIEGLKPLTIYPLLSVITG 449
QY 127 LTEIILPFTFANSYNRKLKNSLNLVALGFGFSPFDVAMGQSSALGFGFELSFKKS 178
DB 450 LLMVYVFNPAAWLNHLLNGLSLGSLNIMLLGLLVIGAMMAIDMGSGPFNKA 501

RESULT 13
US-09-255-829-22
; Sequence 22, Application US/09255829
; Patent No. 6461617
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN, & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/255,829
; FILING DATE: 23-FEB-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB97/02273
; FILING DATE: 22-AUG-1997
; PRIOR APPLICATION DATA:

Query Match **7.9%**; Score **71.5**; DB **4**; Length **856**;
Best Local Similarity **21.5%**; Pred. No. **10**;
Matches **32**; Conservative **30**; Mismatches **56**; Indels **3**

| | |
|--------|--|
| Y D | 48 IVPEFL--LNLFGRGIGSGAQGDILGGSLIIFLD---AVGGLILAGAYLDD- :- :: : 620 IVPVYIGIALNVGNETAKEFNENAFETACIASLTLEPIPELLIPPVVGAFLELEYIDN- ----- 95 KALDGCTTKKAAFOHWNGRGCVMICAGVVWNAVMEPL-----TEFTLL- |
|--------|--|

```

Db      680  K T I D N A L T K R N E K W S D M Y G L I V A Q W L S T V N T Q F T T I K E G M Y K A L N T Q A Q A L E E I I K V R Y - 738
QY      138  N S Y N R K L K N S L N V A L G G F E P S F D V A M G S 166
Db      739  N I Y S E K E K S N I N I D F N D I N S K L N E G I N O A 767

```

NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN, & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/255,829

```

PRIORITY APPLICATION DATA: PCT/GB97/02273
APPLICATION NUMBER: 22-AUG-1997
FILING DATE: 22-AUG-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/782,893
FILING DATE: 27-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: ESMOND, ROBERT W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 1581.0130002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 1169 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-255-829-20

```

Query Match 7.9%; Score 71.5; DB 4; Length 1169;
Best Local Similarity 21.5%; Pred. No. 16;
Matches 32: Conservative 30; Mismatches 56; Indels 31

UNITED STATES DEPARTMENT OF JUSTICE
(USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 20:13:37 ; Search time 29.7835 seconds
(without alignments)
7815.311 Million cell updates/sec

Title: US-09-508-487-18
Perfect score: 759
Sequence: 1 attgttaaagaattgaaat.....attctgtattgaatggggtg 759

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------------------------|
| C 1 | 44.8 | 5.9 | 6243 | 2 | US-09-056-075-1 Sequence 1, Appl |
| C 2 | 44.4 | 5.8 | 2032 | 4 | US-09-241-581B-5 Sequence 5, Appl |
| C 3 | 44.4 | 5.8 | 2032 | 4 | US-08-265-428-5 Sequence 5, Appl |
| C 4 | 44.4 | 5.8 | 2032 | 5 | PCT-US95-07721-5 Sequence 5, Appl |
| C 5 | 43 | 5.7 | 6243 | 2 | US-09-056-075-1 Sequence 1, Appl |
| C 6 | 41.6 | 5.5 | 5852 | 1 | US-07-867-106-2 Sequence 2, Appl |
| C 7 | 41.4 | 5.5 | 2547 | 3 | US-08-262-220-7 Sequence 7, Appl |
| C 8 | 41.4 | 5.5 | 2547 | 3 | US-08-471-733-7 Sequence 7, Appl |
| C 9 | 41.4 | 5.5 | 2547 | 3 | US-08-468-878-7 Sequence 7, Appl |
| C 10 | 41.4 | 5.5 | 2547 | 4 | US-08-750-494-7 Sequence 7, Appl |
| C 11 | 40.8 | 5.4 | 2251 | 4 | US-08-991-677-11 Sequence 11, Appl |
| C 12 | 40.8 | 5.4 | 5733 | 2 | US-08-473-553A-1 Sequence 1, Appl |
| C 13 | 40 | 5.3 | 998 | 4 | US-09-122-400B-5 Sequence 5, Appl |
| C 14 | 39.8 | 5.2 | 796 | 4 | US-08-936-165A-103 Sequence 103, App |
| C 15 | 39.2 | 5.2 | 2973 | 1 | US-08-451-715A-7 Sequence 7, Appl |
| C 16 | 39 | 5.1 | 237 | 3 | US-08-680-726A-59 Sequence 59, Appl |
| C 17 | 39 | 5.1 | 237 | 3 | US-09-092-409-59 Sequence 59, Appl |
| C 18 | 39 | 5.1 | 3077 | 3 | US-09-410-028-3 Sequence 3, Appl |
| C 19 | 39 | 5.1 | 10592 | 1 | US-08-680-726A-51 Sequence 51, Appl |
| C 20 | 39 | 5.1 | 10592 | 1 | US-08-680-726A-52 Sequence 52, Appl |
| C 21 | 39 | 5.1 | 10592 | 3 | US-09-092-409-51 Sequence 51, Appl |
| C 22 | 39 | 5.1 | 10592 | 3 | US-09-092-409-52 Sequence 52, Appl |
| C 23 | 38.8 | 5.1 | 1648 | 3 | US-09-165-241-2 Sequence 2, Appl |
| C 24 | 38.8 | 5.1 | 9636 | 1 | US-08-323-170B-1 Sequence 1, Appl |
| C 25 | 38.8 | 5.1 | 9636 | 4 | US-08-954-441-1 Sequence 1, Appl |
| C 26 | 38.6 | 5.1 | 1850 | 3 | US-08-617-860B-32 Sequence 32, Appl |
| C 27 | 38.6 | 5.1 | 4098 | 2 | US-08-605-106-4 Sequence 4, Appl |

| | | | | | | |
|------|------|-----|-------|---|---------------------|-------------------|
| 28 | 38.6 | 5.1 | 20674 | 4 | US-09-641-638-651 | Sequence 651, App |
| 29 | 38.4 | 5.1 | 1103 | 4 | US-09-122-400B-20 | Sequence 20, Appl |
| 30 | 38.4 | 5.1 | 1639 | 2 | US-08-468-819-77 | Sequence 77, Appl |
| C 31 | 38.2 | 5.0 | 1119 | 4 | US-09-178-973B-7 | Sequence 7, Appl |
| C 32 | 38.2 | 5.0 | 1119 | 4 | US-09-419-568F-7 | Sequence 7, Appl |
| C 33 | 38.2 | 5.0 | 1119 | 4 | US-09-354-243B-7 | Sequence 8, Appl |
| C 34 | 38.2 | 5.0 | 7445 | 4 | US-09-178-973B-8 | Sequence 8, Appl |
| C 35 | 38.2 | 5.0 | 7445 | 4 | US-09-419-568F-8 | Sequence 8, Appl |
| C 36 | 38.2 | 5.0 | 7445 | 4 | US-09-354-243B-8 | Sequence 8, Appl |
| C 37 | 37.8 | 5.0 | 651 | 4 | US-09-134-001C-1749 | Sequence 1749, Ap |
| C 38 | 37.8 | 5.0 | 1599 | 1 | US-08-143-219-27 | Sequence 27, Appl |
| C 39 | 37.8 | 5.0 | 19124 | 2 | US-08-487-826B-13 | Sequence 13, Appl |
| C 40 | 37.6 | 5.0 | 828 | 4 | US-08-998-416-538 | Sequence 538, App |
| C 41 | 37.6 | 5.0 | 1512 | 2 | US-08-370-581-1 | Sequence 1, Appl |
| C 42 | 37.6 | 5.0 | 1512 | 4 | US-09-146-893-1 | Sequence 1, Appl |
| C 43 | 37.6 | 5.0 | 11093 | 2 | US-08-723-306-5 | Sequence 5, Appl |
| C 44 | 37.6 | 5.0 | 11093 | 5 | PCT-US96-10041-5 | Sequence 5, Appl |
| C 45 | 37.4 | 4.9 | 2015 | 3 | US-08-633-993A-30 | Sequence 30, Appl |

ALIGNMENTS

RESULT 1
US-09-056-075-1/c
; Sequence 1, Application US/09056075
; Patent No. 5955368
; GENERAL INFORMATION:
; APPLICANT: Johnson, Eric A.
; APPLICANT: Bradshaw, Marite
; APPLICANT: Rood, Julian
; TITLE OF INVENTION: Expression System for Clostridium
; TITLE OF INVENTION: Species
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/056,075
; APPLICATION NUMBER: 27386
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95238
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6243 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3770..4013
; OTHER INFORMATION: /note= "RP4 origin of DNA transfer (orit) from
; OTHER INFORMATION: plasmid RP4"
; US-09-056-075-1

Query Match 5.9%; Score 44.8; DB 2; Length 6243;
Best Local Similarity 48.7%; Pred. No. 0.2;


```

RESULT 10
US-08-750-494-7/c
; Sequence 7, Application US/08750494
; Patent No. 6204018
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM SVEN
; APPLICANT: BARBOUR ALAN G.
; TITLE OF INVENTION: NEW 56 KDA ANTIGEN FROM BORRELIA
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 SEVENTH STREET, N.W.
; CITY: WASHINGTON
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/750,494
; FILING DATE:
; CLASSIFICATION:

```

; LENGTH: 2251

; LENGTH: 2251

```
; TYPE: DNA
; ORGANISM: Pinus taeda
US-08-991-677-11

Query Match      5.4%; Score 40.8; DB 4; Length 2251;
Best Local Similarity 48.0%; Pred. No. 1.3;
Matches 117; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 1 ATGTGTTAAAGAAATGAAATGATAATTTTATGTCACAAATCAAGAACCTCTATTGGGAAG 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 619 ACTATAAATTCAGTTGGAATGAGTGTGTTATGATGTTTATAGATGTTAAATTTATATA 678

QY 61 CGAATTTCAAGCAATAATTTGAAAAAGTTAAATTTAAATAAATTTAAAAACCTTTTAA 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 679 TGTAAATTAATTTATTTTGAATACAAAAATATATATTTGGATAAAAAATTTGTTGTT 738

QY 121 AATTCATTAATAGTCACCATAGTACAGCTTTTAAATAAAGGGGTTTTTATGAATAAAT 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 739 AAATTTAGAGTAAAAATTTCAAAATCTAAAAATATTAACACACTATTATTATTTAAAAAAT 798

QY 181 TTTAATTTTGTGTTGGCAACCTTTTGTGTTTCTTCTAGCTTTGCTCAAGCTAATGATTC 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 799 TGTGGTAAATTTTATCTTATATTTAAGTTAAAAATTTAGAAAAAATTAATTTAAATTA 858

QY 241 TAAA 244
    ||| |||
Db 859 TAAA 862

RESULT 12
US-08-473-553A-1/c
; Sequence 1, Application US/08473553A
; Patent No. 5859338
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliot M.
; APPLICANT: Clark, Steven E.
; APPLICANT: Williams, Robert W.
; TITLE OF INVENTION: Plant Clavatal Nucleic Acids,
; TITLE OF INVENTION: Transformed Plants, and Proteins
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,553A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-60886/RET/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5733 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: 2434..5037
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5117..5467
US-08-473-553A-1

Query Match      5.4%; Score 40.8; DB 2; Length 5733;
Best Local Similarity 48.3%; Pred. No. 1.5;
Matches 114; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 10 AGAATGAAATGATAATTTTATGTCACAAATCAAGAACCTCTATTGGGAAGCAATTTCA 69
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1153 AAGTCCTCAACCGATCAATTAATAGCAACTGCCGATACCAAAATATATATCGTTTGTAA 1094

QY 70 AGCAATAATTTGAAAAAGTTAAATTTAAATAAATTTAAAAACCTTTTAAATTTCAAT 129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1093 TCGGTGATTCAGAAAAAAGTTCAATGATATATAAATTCGTCGTAGTGTGATATTAGGAC 1034

QY 130 AATATGCTACCATAGTACCAGCTTTTAAATAAAGGGTTTTTATGAATAAACCTTTTAAAT 189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1033 AAATCACTAGTTTGAACATAAGATACATAAAACATATTTAATGTTGAAAGTTTAGGTTT 974

QY 190 TGTTTTGGCAACCTTTTGTGTTTTTCTAGCTTTGCTCAAGCTAATGATTTCTAAA 245
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 973 AGTTGTTAATACTTTTGGCTGAAAAACGTGTCATTTTGTCTTAATGATGCAAAA 918

RESULT 13
US-09-122-400B-5/c
; Sequence 5, Application US/09122400B
; Patent No. 6245974
; GENERAL INFORMATION:
; APPLICANT: Michalowski, Susan
; APPLICANT: Spiker, Steven
; TITLE OF INVENTION: MATRIX ATTACHMENT REGIONS
; FILE REFERENCE: Michalowski and Spiker
; CURRENT APPLICATION NUMBER: US/09/122,400B
; CURRENT FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 60/066,118
; PRIOR FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; TYPE: DNA
; LENGTH: 998
; ORGANISM: Nicotiana tabacum
US-09-122-400B-5

Query Match      5.3%; Score 40; DB 4; Length 998;
Best Local Similarity 46.0%; Pred. No. 1.7;
Matches 172; Conservative 0; Mismatches 200; Indels 2; Gaps 1;

QY 5 TTAAGAAGATTGAAATTCATAATTTTATGGTCAAAATCAAGAAAGCTCTATTGGGAAGCAA 64
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 733 TGAATAATATTTTCTCAATAATAATATTTTGTAAAACTGAAAAAAGCAAGTAAT 674

QY 65 TTTCACGCAATAATTTGAAAAAGTTAAATTTAAATAACTTTAAAAACCTTTTAAAT 124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 673 TTTCTAAAGTAATTTAAAAACTTGAATAAATTTTAACTAAAACTGAAAAA 614

QY 125 TCATTAATATGCTACCATAGTACCAGTTTAAATAAAGGGTTTTTATGAATAAACTTTTA 184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 613 AGAAGAAATATTTTCTTTTAA--GTTTTTACAAAAATATGTTTGAATAATATTTC 556

QY 185 ATTTTGTGTTGGCAACCTTTTGTGTTTTTCTAGCTTTGCTCAAGCTAATGATTTCTAAA 244
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 555 GCTTTTAAAGCAGTTTTTTTGTAAAAATTTAGAAAAAATAATTTTTCATTTTCA 496

QY 245 AATGGTCGGTTTGGGATGAGTGGAGAAAAAATTTTGGTTTTATGAACATAGCAAGCAA 304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 495 GTTTTGTAGTAGTTTTTCCAGATTTTACAAATAAAATGCTTTAGAAAAATTTGATTT 436

QY 305 GATCCTATTGTACCATTATTTATGAACCTTTTATGAGGTTTGGAAATAGGCTCTCTTGCT 364
```

Db 435 CTCTCTTTTTCAGTTTTTACAAAACATCTGTTTAGAAAAATATTTTCAATATTTTCT 376
QY 365 CAAGGAGATATCT 378
Db 375 AAATAGTTTTTT 362

RESULT 14

US-08-936-165A-103/C
; Sequence 103, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
; POLYPEPTIDES AND THEIR USES
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,165A
; FILING DATE: 24-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,032
; FILING DATE: 24-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Glimml, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:

INFORMATION FOR SEQ ID NO: 103:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 796 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-08-936-165A-103

Query Match 5.2%; Score 39.8; DB 4; Length 796;

Best Local Similarity 46.7%; Pred. No. 1.8; Mismatches 122; Indels 0; Gaps 0;

QY 94 TTTAATAATCTTTAAACCTTTTAAATTTTCAATATGCTACCATGACAGTTT 153
Db 684 TTCAATAATAATTAATTAATTTGATATTTCTTTNACGTTACCANAATACATAT 625
QY 154 TAATAAGGGGTTTTATGAATAAACTTTTAAATTTTGTGTTGGCAACCTTTGTGTTT 213
Db 624 ACAATCAATTAATTTCTTAAATAAATTTGACTTATTATTATGTCATTATTTTAT 565

QY 214 TTCTAGCTTTTGCTCAAGCTAATCATCTTAAAAATGTCGTTTGGGATGAGTCTGGAGA 273
Db 564 TATTGTTTCATTTGAGNTAAACCCCTTTTGAATCAAGGTTTTAGTAGGAAAAATCCATA 505
QY 274 AAAACTTTTGGTTTATGAAGACTAGCAAGCAAGATCTATTGTACCACTTT 322
Db 504 AACCCCGGTTTTNAAGANTAATGTCATTTNAGCNAATAATTGTCAGTTT 456

RESULT 15

US-08-451-715A-7/c
; Sequence 7, Application US/08451715A
; Patent No. 5801013
; GENERAL INFORMATION:
; APPLICANT: Tao, Jianshi
; APPLICANT: Qui, Yan
; APPLICANT: Homan, Fariba
; APPLICANT: Shen, Xiaoyu
; APPLICANT: Schimmel, Paul R.
; TITLE OF INVENTION: Helicobacter Aminoacyl-tRNA Synthetase
; PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,715A
; FILING DATE: 26-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: CPI94-25
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2973 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 219..2834
US-08-451-715A-7

Query Match 5.2%; Score 39.2; DB 1; Length 2973;

Best Local Similarity 49.5%; Pred. No. 3;
Matches 101; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 19 ATTGATAATTTTATGGTCAATCAAGAAGCTCTATTGGGAAGGAAATTTCAAGCAATAAT 78
Db 322 ATGTCTCATTCGCCATCAATTTTCAAAATACCCCTTATGAGAGCAAAATTTCAATAATCTTT 263
QY 79 TTGAAAAAGTTTAAATTTTAAATAACTTTTAAAAACCTTTTAAATTTCAATTAATGCTA 138
Db 262 TTTTCTATCTCTTCGTTGGTAGGGGTTCTTGTTCATTATTATTATTAATCTTA 203
QY 139 CCATGATACCAAGTTTATAAAGGGGTTTTATGAATAAACTTTTAAATTTTGTGTTTGGC 198
Db 202 AAATAGAGCGTCTTTTCAAAAAATGTTGAATTTTGAACGCCTTTTGTGCTTTTACGCTTTTA 143

Oy 199 AACCTTTTGTGTTTTTCTAGCTT 222
| | | | | | | | | | | | | | | | | |
Db 142 ATTGTTGGGTATTTTTTCAAAATT 119

Search completed: March 15, 2003, 07:43:26
Job time : 60.7835 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 22:48:52 ; Search time 49.6392 Seconds
(without alignments)

10731.744 Million cell updates/sec

Title: US-09-508-487-18

Perfect score: 759

Sequence: 1 attgttaagaagaattgaat.....attctgtattgaaatgggtg 759

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 501302 seqs, 350932545 residues

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA.*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 51.2 | 6.7 | 416 | 10 | US-09-960-352-4584 |
| 2 | 46 | 6.1 | 640681 | 10 | US-09-790-988-1 |
| 3 | 45.6 | 6.0 | 2000 | 9 | US-09-938-842A-5061 |
| 4 | 44.2 | 5.8 | 513509 | 9 | US-09-754-853A-4 |
| 5 | 43 | 5.7 | 3979 | 9 | US-09-989-919-52 |
| 6 | 42.2 | 5.6 | 419 | 10 | US-09-960-352-11234 |
| 7 | 42 | 5.5 | 5520 | 9 | US-10-001-887-43 |
| c 8 | 41.8 | 5.5 | 337 | 10 | US-09-960-352-6036 |
| 9 | 41.4 | 5.5 | 2846 | 9 | US-10-008-016-1 |
| c 10 | 41 | 5.4 | 640681 | 10 | US-09-790-988-1 |
| 11 | 40.8 | 5.4 | 431 | 10 | US-09-969-373-739 |
| 12 | 40.8 | 5.4 | 2251 | 10 | US-09-796-256A-11 |
| 13 | 40.8 | 5.4 | 127197 | 9 | US-09-754-853A-1 |
| 14 | 40.6 | 5.3 | 352 | 10 | US-09-960-352-10129 |
| 15 | 40.6 | 5.3 | 414 | 10 | US-09-960-352-6528 |
| c 16 | 40.6 | 5.3 | 2846 | 9 | US-10-008-016-1 |
| 17 | 40.6 | 5.3 | 20966 | 10 | US-09-776-976-7 |
| 18 | 40.6 | 5.3 | 20966 | 10 | US-09-758-055-7 |
| 19 | 40.6 | 5.3 | 20966 | 10 | US-09-909-547-7 |

| | | | | | | |
|------|------|-----|-------|----|---------------------|-------------------|
| c 20 | 40.4 | 5.3 | 416 | 10 | US-09-960-352-4584 | Sequence 4584, Ap |
| 21 | 40.2 | 5.3 | 230 | 10 | US-09-969-373-1411 | Sequence 1411, Ap |
| c 22 | 40.2 | 5.3 | 335 | 10 | US-09-960-352-13489 | Sequence 13489, A |
| c 23 | 40.2 | 5.3 | 73308 | 10 | US-09-954-456-2276 | Sequence 2276, Ap |
| c 24 | 40 | 5.3 | 988 | 10 | US-09-867-550-1405 | Sequence 1405, Ap |
| c 25 | 40 | 5.3 | 998 | 10 | US-09-816-894-5 | Sequence 5, Appli |
| c 26 | 39.8 | 5.2 | 796 | 10 | US-09-939-980-103 | Sequence 103, App |
| c 27 | 39.8 | 5.2 | 1674 | 9 | US-09-938-842A-3489 | Sequence 3489, Ap |
| c 28 | 39.8 | 5.2 | 3935 | 10 | US-09-925-297-141 | Sequence 141, App |
| c 29 | 39.4 | 5.2 | 411 | 10 | US-09-960-352-14521 | Sequence 14521, A |
| 30 | 39.4 | 5.2 | 2000 | 9 | US-09-938-842A-4370 | Sequence 4370, Ap |
| c 31 | 39 | 5.1 | 5690 | 10 | US-09-773-416-13 | Sequence 13, Appl |
| c 32 | 39 | 5.1 | 2699 | 8 | US-09-764-176-9 | Sequence 9, Appli |
| 33 | 38.8 | 5.1 | 2798 | 8 | US-08-834-666A-3 | Sequence 3, Appli |
| 34 | 38.8 | 5.1 | 2798 | 8 | US-08-834-666A-1 | Sequence 1, Appli |
| c 35 | 38.6 | 5.1 | 263 | 10 | US-09-960-352-7661 | Sequence 7661, Ap |
| c 36 | 38.6 | 5.1 | 371 | 10 | US-09-969-373-1447 | Sequence 1447, Ap |
| c 37 | 38.6 | 5.1 | 981 | 10 | US-09-974-300-4908 | Sequence 4908, Ap |
| c 38 | 38.6 | 5.1 | 1008 | 10 | US-09-974-300-4909 | Sequence 4909, Ap |
| c 39 | 38.6 | 5.1 | 2000 | 9 | US-09-938-842A-4157 | Sequence 4157, Ap |
| c 40 | 38.6 | 5.1 | 2000 | 9 | US-09-938-842A-4784 | Sequence 4784, Ap |
| c 41 | 38.6 | 5.1 | 2008 | 9 | US-10-081-051-84 | Sequence 84, Appl |
| 42 | 38.6 | 5.1 | 2172 | 9 | US-09-966-880A-15 | Sequence 15, Appl |
| 43 | 38.6 | 5.1 | 2818 | 9 | US-09-966-880A-7 | Sequence 7, Appli |
| 44 | 38.6 | 5.1 | 6564 | 9 | US-09-966-880A-10 | Sequence 10, Appl |
| 45 | 38.6 | 5.1 | 11204 | 9 | US-09-966-880A-35 | Sequence 35, Appl |

ALIGNMENTS

RESULT 1

US-09-960-352-4584

; Sequence 4584, Application US/09960352

; Patent No. US20020137139A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Tao, Nengbing

; APPLICANT: Wyatt, John C.

; APPLICANT: Mathialagan, Nagappan

; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION A

; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION

; FILE REFERENCE: 16511.006/37-21(10298)C

; CURRENT APPLICATION NUMBER: US/09/960,352

; CURRENT FILING DATE: 2001-09-24

; NUMBER OF SEQ ID NOS: 15112

; SEQ ID NO 4584

; LENGTH: 416

; TYPE: DNA

; ORGANISM: Bos taurus

; OTHER INFORMATION: Clone ID: 20-LIB3057-016-Q1-K1-E11

US-09-960-352-4584

Query Match

Best Local Similarity 6.7%; Score 51.2; DB 10; Length 416;

Matches 122; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 7 AAAGAATTGAATTGATTAATTTTATGGTCAATCAATCAAGAGCTCTATTGGGAGCGAAT 66

Db 164 AAAAAAAAAAAAAAAAAAGAGATTTTAAATAATTAATAATAATAATAATAATTCAT 223

QY 67 TCAAGCAATAATTGCAAAAAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 126

Db 224 AAAAAAAAAATTTTAAAGAAAAATTTTAAATTTTAAAGAAAAATTTTAAAGAAAAAT 283

QY 127 ATTAATATCTACCTAGTACCAGTTTAAATAAGGGTTTTATGAATAAACTTTAAAT 186

Db 284 TTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 343

QY 187 TTTTCTTTGGCAACCTTTTGTGTTTCTAGCTTGTCTCAAGCTTAATGATTTCTAAAAA 246

Db 344 AATAATATGGTAAATTTTATAAATTAATAAATTTTATAAATAAATAAATTTTAAATTAATAA 403

| Db | 619 | ACATATAAATTGAGTTGGAATGAGTGTTTTATGGATTTTTTAAGATGTTTAAATTTATATA | 678 |
|----|-----|--|-----|
| QY | 61 | CGAATTTCAAGCAATAATTTGAAAAAAGCTTAAATTTAAATAACTTTTAAAAACCTTTTTTA | 120 |
| Db | 679 | TGTAATTTAAAAATTTTATTTTGAATAACAAAAATATATAATTCGATAAAAAATGTTTGT | 738 |
| QY | 121 | AATTCATTAATATGCTACCATAGTACAGGTTTTTAATAAAGGGGTTTTTATGAATAAAT | 180 |
| Db | 739 | AAATTTAGAGTAAAAATTTTCAAAATCTAAAAATAAATAAACAATATATTTTAAAAAAT | 798 |
| QY | 181 | TTTAATTTTGTGTTTGCAACCTTTTGTGTTTTTTCTAGCTTGTCTCAAGCTAATGATTC | 240 |
| Db | 799 | TGTTGGTAAATTTTATCTTATATTAAGCTTAAATTTAGAAAAATTAATTTTAAATTA | 858 |
| QY | 241 | TAAA 244 | |
| Db | 859 | TAAA 862 | |

RESULT 13

US-09-754-853A-1

; Sequence 1, Application US/09754853A

; Publication No. US20030005491A1

; GENERAL INFORMATION:

; APPLICANT: Hauge, Brian M.

; APPLICANT: Parnell, Laurence D.

; APPLICANT: Parsons, Jeremy D.

; APPLICANT: Wang, Ming Li

; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With

; FILE REFERENCE: Soybean Cyst Nematode Resistance

; CURRENT FILING DATE: 2001-01-05

; PRIOR APPLICATION NUMBER: US 60/174,880

; PRIOR FILING DATE: 2000-01-07

; NUMBER OF SEQ ID NOS: 1119

; SEQ ID NO 1

; LENGTH: 127197

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: 515002_region_G2

US-09-754-853A-1

Query Match 5.4%; Score 40.8; DB 9; Length 127197;

Best Local Similarity 48.3%; Pred. No. 38;

Matches 114; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 89 TTAATTTAAATAACTTTTAAAAACCTTTTAAATTTTCATTAATATGCTTACCATAGTACC 148

Db 71508 TGATATATATATATATAATCATATCGTATAACCTTCATACTATATATATTTTAA 71567

QY 149 AGTTTTTAATAAGGGTTTTTATGAATAAATTTTAAATTTTGTGTTTGGCAACCTTTGT 208

Db 71568 ACTTTTAATTAATATTTAATATTTTCTTAAAAATTTATCATTTAAAAATTCATTTT 71627

QY 209 GTTTTTTCTAGCTTGTCTCAAGCTAATCATCTTAAAAATGGTCGTTGGGATGAGTCT 268

Db 71628 ATATAATATATATACAAAATCTAATAATTAATCTAAAAATTAATTTTGTTTATA 71687

QY 269 GGAGAAAAACCTTTTGGTTTATGAACACTAGCAAGAGTCCCTATTGTACCATTTT 324

Db 71688 TAGATTAAAAATTGATCATATATGAATATTTTAAAAATATAAATATATGATATTTT 71743

RESULT 14

US-09-960-352-10129

; Sequence 10129, Application US/09960352

; Patent No. US20020137139A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Tao, Nengbing

; APPLICANT: Byatt, John C.

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:19:56 ; Search time 31.1742 Seconds
(without alignments)
765.116 Million cell updates/sec

Title: US-09-508-487-19

Perfect score: 902

Sequence: 1 MNKLIFVLATFCVFSFAQ.....DVAMQSSALGFLSFKKSY 179

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 902 | 100.0 | 179 | 20 AAY19808 | B. burgdorferi ant |
| 2 | 902 | 100.0 | 179 | 20 AAY04278 | Borrelia burgdorfe |
| 3 | 799 | 88.6 | 158 | 20 AAY19809 | B. burgdorferi ant |
| 4 | 769.5 | 85.3 | 178 | 20 AAY04279 | Borrelia afzelii A |
| 5 | 737 | 81.7 | 177 | 20 AAY04280 | Borrelia garinii I |
| 6 | 155.5 | 17.2 | 95 | 20 AAY04282 | p11 protein sequen |
| 7 | 128 | 14.2 | 25 | 20 AAY04281 | Borrelia burgdorfe |
| 8 | 84.5 | 9.4 | 378 | 22 AAB88515 | Haemophilus influe |
| 9 | 84.5 | 9.4 | 378 | 23 AAU91446 | Haemophilus influe |
| 10 | 84 | 9.3 | 355 | 23 ABB93914 | Herbicidally activ |

| | | | | | | |
|----|------|-----|------|----|----------|--------------------|
| 11 | 84 | 9.3 | 1456 | 22 | ABG07111 | Novel human diagno |
| 12 | 84 | 9.3 | 1456 | 22 | ABG24514 | Novel human diagno |
| 13 | 83.5 | 9.3 | 384 | 22 | AAG89852 | C glutamicum prote |
| 14 | 83.5 | 9.3 | 384 | 22 | AAB78951 | C. glutamicum SRT |
| 15 | 82 | 9.1 | 369 | 22 | AAB78952 | C. glutamicum SRT |
| 16 | 82 | 9.1 | 553 | 20 | AAY01651 | A protein with cat |
| 17 | 82 | 9.1 | 553 | 21 | AAB20579 | Mouse OCTN1 amino |
| 18 | 81.5 | 9.0 | 572 | 21 | AAY81595 | Streptococcus pneu |
| 19 | 81 | 9.0 | 325 | 23 | ABP25872 | Streptococcus poly |
| 20 | 81 | 9.0 | 504 | 23 | ABP27031 | Streptococcus poly |
| 21 | 80.5 | 8.9 | 111 | 23 | ABP09270 | Human ORFX protein |
| 22 | 80.5 | 8.9 | 650 | 22 | AAU37748 | Streptococcus pneu |
| 23 | 80 | 8.9 | 243 | 23 | ABP38116 | Staphylococcus epi |
| 24 | 80 | 8.9 | 487 | 22 | AAG91121 | C glutamicum prote |
| 25 | 80 | 8.9 | 487 | 22 | AAB78989 | C. glutamicum SRT |
| 26 | 79.5 | 8.8 | 319 | 22 | AAU38331 | Salmonella typhi c |
| 27 | 79.5 | 8.8 | 345 | 22 | AAB76781 | Corynebacterium gl |
| 28 | 79.5 | 8.8 | 417 | 22 | AAG93096 | C glutamicum prote |
| 29 | 79 | 8.8 | 324 | 22 | AAB76751 | Corynebacterium gl |
| 30 | 79 | 8.8 | 346 | 22 | AAG92057 | C glutamicum prote |
| 31 | 79 | 8.8 | 346 | 22 | AAB76750 | Corynebacterium gl |
| 32 | 78.5 | 8.7 | 513 | 22 | AAG93106 | C glutamicum prote |
| 33 | 77.5 | 8.6 | 392 | 20 | AAU00139 | Enterococcus faeca |
| 34 | 77.5 | 8.6 | 392 | 23 | ABP43358 | E faecalis EF068 a |
| 35 | 77.5 | 8.6 | 502 | 20 | AAW99600 | E stationis multi- |
| 36 | 77.5 | 8.6 | 513 | 23 | AAU97210 | Corn Beta vulgaris |
| 37 | 77.5 | 8.6 | 1638 | 20 | AAU00138 | Enterococcus faeca |
| 38 | 77.5 | 8.6 | 1638 | 20 | AAU00140 | Enterococcus faeca |
| 39 | 77.5 | 8.6 | 1638 | 20 | AAU00142 | Enterococcus faeca |
| 40 | 77.5 | 8.6 | 1638 | 23 | ABP43357 | E faecalis EF068 p |
| 41 | 77.5 | 8.6 | 1638 | 23 | ABP43359 | E faecalis EF069 p |
| 42 | 77.5 | 8.6 | 1638 | 23 | ABP43361 | E faecalis EF070 p |
| 43 | 77 | 8.5 | 251 | 22 | AAU04855 | Micromonospora eve |
| 44 | 76.5 | 8.5 | 1755 | 21 | AAG53006 | Arabidopsis thalia |
| 45 | 76.5 | 8.5 | 1793 | 21 | AAG53005 | Arabidopsis thalia |

ALIGNMENTS

RESULT 1

AAU19808
ID AAY19808 standard; Protein; 179 AA.

XX AAY19808;

XX 19-JUL-1999 (first entry)

XX B. burgdorferi antigenic protein, f933.aa.

XX Antigenic protein; vaccine; Lyme disease; infection; detection.

XX Borrelia burgdorferi.

XX WO9859071-A1.

XX 30-DEC-1998.

XX 18-JUN-1998; 98WO-US12718.

XX 03-SEP-1997; 97US-0057483.

XX 20-JUN-1997; 97US-0050359.

XX 22-JUL-1997; 97US-0053344.

XX 22-JUL-1997; 97US-0053377.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (MEDI-) MEDIMUNE INC.

XX Choi GH, Erwin AL, Hanson MS, Lathigra R;

XX WPI; 1999-189980/16.

XX N-PSDB; AAX61505.

AAU91446

XX DT 31-MAY-2002 (first entry)
XX DE Herbicidally active polypeptide SEQ ID NO 3125.
XX KW Herbicidal; plant; agriculture; herbicide.
XX OS Arabidopsis thaliana.
XX PN WO200210210-A2.
XX PD 07-FEB-2002.
XX PF 28-AUG-2001; 2001WO-EP09892.
XX PR 28-AUG-2001; 2001WO-EP09892.
XX RA (FARB) BAYER AG.
XX TI Tietjen K, Weidler M;
XX WI WPI; 2002-269010/31.
XX PS Identifying plant target proteins for herbicidally active compounds,
XX PT comprising aligning and comparing nucleic acid or amino acid sequences
XX PT from plant with nucleic acid or amino acid sequences from non-plant
XX PT organisms -
XX PS Claim 5; SEQ ID NO 3125; 261pp + Sequence Listing; English.
XX CC The invention relates to identifying target proteins
XX CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
XX CC aligning and comparing nucleic acid or amino acid sequences from plant
XX CC with nucleic acid or amino acid sequences from non-plant organisms using
XX CC suitable search parameters, where plant sequences having an E-value
XX CC greater by a factor of 3 than the E-value of most similar non-plant
XX CC sequences are selected. The polypeptides or nucleic acids encoding them
XX CC are useful for identifying modulators. The identified modulators are
XX CC useful as herbicides.
XX SQ Sequence 355 AA;
Query Match 9.3%; Score 84; DB 23; Length 355;
Best Local Similarity 29.6%; Pred. No. 0.54;
Matches 48; Conservative 23; Mismatches 71; Indels 20; Gaps 9;
Db 6 IFVLATFCVPSFAQANDSKNGAFKMSAGEK--LLVYETSKQDPIVPFLLNLFLG-FGIG 62
13 VLCLVIFPLLVSAAEENOCGSGKGAFAKASALKYK-----IIAFTSILTAGVFGVC 66
QY 63 SFAQGDILGSLILGFDVAVGIGLILAGAYLDI--KALDGITKKA-AFQWTWKGVMVLAGV 119
Db 67 LPITFGLKTESNFFMYVKAFAAGVILATGFVHILPDRATESLTSCLEEPPWGDFFPM-TGL 125
QY 120 VTMAVTRLTEIILPFTFANSY-NR----KLKNSLNVALGGFE 156
Db 126 VMAAASILTMLI--ESFASGYLNRSRLAKEGKTLPVSTGEE 165
RESULT 11
ABG07111
ID ABG07111 standard; Protein; 1456 AA.
XX AC ABG07111;
XX DT 13-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #7102.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.

XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX WI WPI; 2001-639362/73.
XX DR N-PSDB; AAS71298.
XX PS New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -
XX PS Claim 20; SEQ ID No 37470; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving or
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG0010-ABG30377 represent novel human
XX CC diagnostic amino acid sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 1456 AA;
Query Match 9.3%; Score 84; DB 22; Length 1456;
Best Local Similarity 21.1%; Pred. No. 3.5;
Matches 43; Conservative 32; Mismatches 75; Indels 54; Gaps 8;
QY 13 CVFSSFAQANDSKNGAFKMSAGEKLLVYETSKQDPI--VPFLLNLFL----- 57
Db 658 CLIESTRNSMDATNKAMLESVAEEM---SVSDGVMRLPLFLAMILPVOLGAATADACT 713
QY 58 -----GFGIGSFAQGDILGSLILGFDVAVGIGLILAGAYLDIKALDG 99
Db 714 FIPVTRDQSDIYEVFNAGSSFGSYAAGD-----VLDMSQGVYVSOLRRYVLVASSDG 767
QY 100 ITKKAFAQW--WGKGVML-AGVVTMAVTRLTEII-----LPFTFANSYNRKL--KNSL 148
Db 768 TSKTATFKMEDFEGQNPVKRGTNIYVNRKISKVVDNGSGSLHSTNAGEQITVCSL 827
QY 149 NVALGGFEPSFDVAMQSSALGFE 172
Db 828 NYNIGQIALSFSKAPDKGTEIAIE 851
RESULT 12
ABG24514
ID ABG24514 standard; Protein; 1456 AA.
XX

AC ABG24514;
XX 18-FEB-2002 (first entry)
XX Novel human diagnostic protein #24505.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR N-PSDB; AAS88701.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 20; SEQ ID No 54873; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1456 AA;
Query Match 9.3%; Score 84; DB 22; Length 1456;
Best Local Similarity 21.1%; Pred. No. 3.5;
Matches 43; Conservative 32; Mismatches 75; Indels 54; Gaps 8;
Oy 13 CVFSSFAQDKNKAGFCMSAGEKLLVYETSKQDPI--VPFLNLFL----- 57
Db 658 CLIESTRNMDATNKAMLESVAENM-----SVSDGVMLPLFLAMILPVQIGATADACT 713
Oy 58 -----GFGIGFQAQDILGGSLILGDFAGVIGLILAGAYLDIKALDG 99
Db 714 FIPVTRQSDIYEVFNAGSFGSVAAGD-----VLDMSQGVYVQLRRYVLSVSDG 767
Oy 100 ITKKAAFOWT--WKGKVM--AGVVTMAVTRTEIF-----LPFTFANSYRKL--KNSL 148
Db 768 TSKTATFKMEDFEGQVPIRKGRNIIYVNRKISVVDNGSGSLLHSFTNAAGEQIVTCSL 827

Oy 149 NVALGGFEPFSDVAMQSSALGFE 172
Db 828 NYNIGQIALSFSPKAPDKGTEIAIE 851
RESULT 13
AAG89852
ID AAG89852 standard; Protein; 384 AA.
XX AAG89852;
AC AAG89852;
XX 26-SEP-2001 (first entry)
DT C glutamicum protein fragment SEQ ID NO: 3606.
DE Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
XX Corynebacterium glutamicum.
OS EPI108790-A2.
XX EPI108790-A2.
PN 20-JUN-2001.
XX 18-DEC-2000; 2000EP-0127688.
PF 16-DEC-1999; 99JP-0377484.
PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX WPI; 2001-376931/40.
DR N-PSDB; AAH65071.
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
XX
PS Claim 17; SEQ ID NO: 3606; 246pp + Sequence Listing; English.
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from Coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 384 AA;
Query Match 9.3%; Score 83.5; DB 22; Length 384;
Best Local Similarity 24.3%; Pred. No. 0.68;
Matches 56; Conservative 24; Mismatches 59; Indels 91; Gaps 12;
Oy 6 IFVLATFCVSSFAQDNDK-----NGAFGMSA----- 33
Db 9 VFVIGVITLLTAGWAANHFASVILIRBQLDVSSVLVNGAFGIYALGLPSLAGGLAD 68
Oy 34 --GEKLLVYETSKQDPIVPFLNLFLGFIGSFAQDILGGSLILGDFAGVIGLILAGAY 91
Db 69 RFGARMVLTGG-----VLSALGNLSL-----LAFHD--GPSLLVGRFIVGLGV----- 110

Search completed: March 14, 2003, 14:24:22
Job time : 33.1742 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:22:16 ; Search time 12.4026 seconds
(without alignments)
1387.454 Million cell updates/sec

Title: US-09-508-487-19
Perfect score: 902
Sequence: 1 MNKLLIFVLATFCVFSFAQ.....DVAMGQSSALGFELSFKKSY 179

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 9613422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|----------|----------------------|
| 1 | 902 | 100.0 | 179 | 2 B70104 | hypothetical prote |
| 2 | 291 | 32.3 | 161 | 2 A70207 | conserved hypotet |
| 3 | 279.5 | 31.0 | 170 | 2 F70243 | conserved hypotet |
| 4 | 235 | 26.1 | 190 | 2 G70239 | conserved hypotet |
| 5 | 91 | 10.1 | 616 | 2 G98325 | probable c4-dicarb |
| 6 | 91 | 10.1 | 616 | 2 AG2957 | hypothetical prote |
| 7 | 87 | 9.6 | 449 | 2 S37867 | hypothetical prote |
| 8 | 86 | 9.5 | 214 | 2 S76338 | hypothetical prote |
| 9 | 86 | 9.5 | 307 | 2 T44611 | cytochrome caa3 ox |
| 10 | 86 | 9.5 | 429 | 2 A10485 | c4-dicarboxylate t |
| 11 | 85.5 | 9.5 | 463 | 2 H69822 | sodium-glutamate s |
| 12 | 85 | 9.4 | 146 | 2 S61393 | hypothetical prote |
| 13 | 85 | 9.4 | 146 | 2 F83814 | Na+/H+ antiporter |
| 14 | 84.5 | 9.4 | 378 | 2 C64142 | unclassified hypotet |
| 15 | 84 | 9.3 | 357 | 2 H70346 | undecaprenyl-phosp |
| 16 | 83.5 | 9.3 | 484 | 2 T34251 | hypothetical prote |
| 17 | 82.5 | 9.1 | 265 | 2 JC4509 | flagellar protein |
| 18 | 81.5 | 9.0 | 174 | 2 T11491 | NADH2 dehydrogenas |
| 19 | 81.5 | 9.0 | 336 | 1 S75272 | cytochrome d ubiqu |
| 20 | 81 | 9.0 | 279 | 2 T35791 | probable transmemb |
| 21 | 81 | 9.0 | 403 | 2 C33958 | hypothetical prote |
| 22 | 81 | 9.0 | 508 | 2 D81325 | probable phosphate |
| 23 | 80.5 | 8.9 | 441 | 2 C96032 | C4-dicarboxylate t |
| 24 | 80.5 | 8.9 | 453 | 2 A33597 | C4-dicarboxylate t |
| 25 | 80.5 | 8.9 | 650 | 2 C95101 | PTS system, fructo |
| 26 | 80.5 | 8.9 | 650 | 2 D97969 | phosphotransferase |
| 27 | 80.5 | 8.9 | 666 | 2 A87577 | oligopeptide trans |
| 28 | 80.5 | 8.9 | 709 | 2 S75212 | comE protein - Syn |
| 29 | 80 | 8.9 | 264 | 2 S73384 | hypothetical prote |

| | | | | | | |
|----|------|-----|------|---|--------|--------------------|
| 30 | 80 | 8.9 | 385 | 2 | G71246 | hypothetical prote |
| 31 | 80 | 8.9 | 538 | 2 | A41901 | metabolite export |
| 32 | 80 | 8.9 | 866 | 2 | C97662 | hypothetical prote |
| 33 | 80 | 8.9 | 866 | 2 | AF2886 | conserved hypotet |
| 34 | 79.5 | 8.8 | 319 | 2 | AC0956 | phosphate transpor |
| 35 | 79.5 | 8.8 | 493 | 2 | S77612 | amidophosphoribosy |
| 36 | 79.5 | 8.8 | 692 | 2 | H70362 | glycogen phosphory |
| 37 | 79 | 8.8 | 1090 | 2 | S59077 | cellulose 1,4-beta |
| 38 | 78.5 | 8.7 | 375 | 2 | AH2781 | alcohol dehydrogen |
| 39 | 78.5 | 8.7 | 375 | 2 | H97560 | alcohol dehydrogen |
| 40 | 78.5 | 8.7 | 429 | 2 | C72408 | phosphoglucomutase |
| 41 | 78.5 | 8.7 | 491 | 2 | A70681 | probable dcta prot |
| 42 | 78 | 8.6 | 188 | 2 | E71157 | probable CDP-alcoh |
| 43 | 78 | 8.6 | 235 | 2 | E89938 | conserved hypotet |
| 44 | 78 | 8.6 | 240 | 2 | T38152 | conserved hypotet |
| 45 | 78 | 8.6 | 457 | 2 | AH0898 | PTS system, galact |

ALIGNMENTS

RESULT 1

B70104
hypothetical protein BB0034 - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Mar-2000
C:Accession: B70104
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; W
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; B
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: B70104
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-179 <KLE>
A:Cross-references: GB:AE001117; GB:AE000783; NID:g2687907; PIDN:AAC66426.1; PID:g22
A:Experimental source: strain B31
C:Superfamily: Lyme disease spirochete plasmid hypothetical protein BBA01

| | | | | |
|-----------------------|-----------------|---|-----------|-------------|
| Query Match | 100.0% | Score 902; | DB 2; | Length 179; |
| Best Local Similarity | 100.0%; | Pred. No. 3.5e-73; | | |
| Matches 179; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| QY | 1 | MNKLIFVLATFCVFSFAQANDSKNGAFGMSAGEKLLVYETSKQDPIVPFLLNLFGLFG 60 | | |
| Db | 1 | MNKLIFVLATFCVFSFAQANDSKNGAFGMSAGEKLLVYETSKQDPIVPFLLNLFGLFG 60 | | |
| QY | 61 | IGSFAQGDILGGSLLILGFDVAVGIGLILAGAYLDIKALDGIITKKAFTWKGVMAGV 120 | | |
| Db | 61 | IGSFAQGDILGGSLLILGFDVAVGIGLILAGAYLDIKALDGIITKKAFTWKGVMAGV 120 | | |
| QY | 121 | TMAVTRTEILPTTFANSYNRKLNLSVALGFGFSPFDVAMGQSSALGFELSFKKSY 179 | | |
| Db | 121 | TMAVTRTEILPTTFANSYNRKLNLSVALGFGFSPFDVAMGQSSALGFELSFKKSY 179 | | |

RESULT 2

A70207
conserved hypothetical protein BBA01 - Lyme disease spirochete plasmid A/lp54
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Mar-2000
C:Accession: A70207
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; W
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; B
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: A70207

A: Cross-references: EMBL:AJ010111; NID:94584147; PIDN:CAB40606.1; PID:94584150
A: Experimental source: ATCC 10987
C: Genetix: ctab
C: Superfamily: heme O synthase

Query Match 9.5%; Score 86; DB 2; Length 307;
Best Local Similarity 25.5%; Pred. No. 2.7;
Matches 38; Conservative 27; Mismatches 58; Indels 26; Gaps 7;
QY 53 LNLFLGFGIGSFAQG-DILGSLILGFDVAGIGLILAGA-----YL--DIKALDGIK-- 102
Db 40 LVVTFGLALHFNGLSVNDLKLFFITVIGSLVAGVCCNNYIDRDIDLMERTKTR 99
QY 103 -----KAAFQWTGKGVMLAGVVTMAVTRLTEIL-----PFTFANSYN--RKLNLSLN 149
Db 100 PVTGKYPGFALTEGLVILLGLFVLLTTPMAVLMFGAFTVIVLYSLWTKKRYTLN 159
QY 150 VALGGFEPFDVAMQSS---ALGFELSF 175
Db 160 TVVGSIGAVPLIGWAAIDPSLGHPIAW 188

RESULT 10
A: Cross-references: GB:AL590842; PIDN:CAC93453.1; PID:G15981900; GSPDB:GN00175
A: Experimental source: ATCC 10987
C: Genetix: ctab
C: Superfamily: C4-dicarboxylate carrier protein
Query Match 9.5%; Score 86; DB 2; Length 429;
Best Local Similarity 25.0%; Pred. No. 3.9;
Matches 49; Conservative 30; Mismatches 63; Indels 54; Gaps 10;
QY 4 LLIF-VLATFCVFSFAQANDSKNGAFGMS-----AGEKLLVYETSKQDPIVFFLLNL 55
Db 78 LLYFEIVSTLALLIGLVVNVVAPGV-GMNIDPATLDAKAVALYEAQSQGGIIPFLDI 136
QY 56 FLGFGIGSFAQGDILG---GSLILGFDVAGIG-----LILAGAYLDIKALD 98
Db 137 IPGSVVGAFAAGNQLVLLFAVLFGALHRLGKGLIFNVIESFSRVPFGVNMIMRLA 196
QY 99 GITYKKAQFTWCK-GVMLAGVVTMAVTRLTEILPF-----TFA--NSYN- 141
Db 197 PLGAFGAMAFITIKYGV-----GSLVOLQLILCYLTCILFVVLVLTGTAIKFNGFNI 249
QY 142 ----RKLNLSNLVALG 153
Db 250 FFEIRYKEELLIVLG 265

RESULT 11
A: Cross-references: EMBL:AJ010111; NID:94584147; PIDN:CAB40606.1; PID:94584150
A: Experimental source: ATCC 10987
C: Genetix: ctab
C: Superfamily: heme O synthase
Query Match 9.5%; Score 86; DB 2; Length 307;
Best Local Similarity 25.5%; Pred. No. 2.7;
Matches 38; Conservative 27; Mismatches 58; Indels 26; Gaps 7;
QY 53 LNLFLGFGIGSFAQG-DILGSLILGFDVAGIGLILAGA-----YL--DIKALDGIK-- 102
Db 40 LVVTFGLALHFNGLSVNDLKLFFITVIGSLVAGVCCNNYIDRDIDLMERTKTR 99
QY 103 -----KAAFQWTGKGVMLAGVVTMAVTRLTEIL-----PFTFANSYN--RKLNLSLN 149
Db 100 PVTGKYPGFALTEGLVILLGLFVLLTTPMAVLMFGAFTVIVLYSLWTKKRYTLN 159
QY 150 VALGGFEPFDVAMQSS---ALGFELSF 175
Db 160 TVVGSIGAVPLIGWAAIDPSLGHPIAW 188

A: Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari
Nature 390, 249-256, 1997
A: Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; G
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo,
Kocher, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardi
A: Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; M
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Port
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Sca
A: Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.;
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchi
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshi
A: Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A: Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*
A: Reference number: A69580; MUID:98044033; PMID:9384377
A: Accession: H69822
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-463 <KUN>
A: Cross-references: GB:Z99108; GB:AL009126; NID:92633055; PIDN:CAB12741.1; PID:92633
A: Experimental source: strain 168
C: Genetix: yhcL
C: Superfamily: *Bacillus subtilis* sodium-glutamate symporter homolog yhcL

Query Match 9.5%; Score 85.5; DB 2; Length 463;
Best Local Similarity 24.3%; Pred. No. 4.6;
Matches 28; Conservative 25; Mismatches 43; Indels 19; Gaps 4;
QY 27 GAFGMSAGEK-----LLVYETSKQDPIV-PFLNLFGLFGIGSFAQGDILG-- 72
Db 334 GSGFLSIGQNGCAGIYPAAMLAMTPTVGQNFDFVFIITVIANVAISGFGVAGVGGAAT 393
QY 73 -SLILGFDVAGIGLILAGAYLDIKALDITKKAQFQWTGKGVMLAGVVTMAVTR 126
Db 394 FAALLVLSLNNMPVALAGLLISIEPLIDMGRTALNV-----SGSNTSGILTSKYK 444

RESULT 12
S61393
A: Cross-references: GB:AL590842; PIDN:CAC93453.1; PID:G15981900; GSPDB:GN00175
A: Experimental source: ATCC 10987
C: Genetix: ctab
C: Superfamily: C4-dicarboxylate carrier protein
Query Match 9.5%; Score 86; DB 2; Length 429;
Best Local Similarity 25.0%; Pred. No. 3.9;
Matches 49; Conservative 30; Mismatches 63; Indels 54; Gaps 10;
QY 4 LLIF-VLATFCVFSFAQANDSKNGAFGMS-----AGEKLLVYETSKQDPIVFFLLNL 55
Db 78 LLYFEIVSTLALLIGLVVNVVAPGV-GMNIDPATLDAKAVALYEAQSQGGIIPFLDI 136
QY 56 FLGFGIGSFAQGDILG---GSLILGFDVAGIG-----LILAGAYLDIKALD 98
Db 137 IPGSVVGAFAAGNQLVLLFAVLFGALHRLGKGLIFNVIESFSRVPFGVNMIMRLA 196
QY 99 GITYKKAQFTWCK-GVMLAGVVTMAVTRLTEILPF-----TFA--NSYN- 141
Db 197 PLGAFGAMAFITIKYGV-----GSLVOLQLILCYLTCILFVVLVLTGTAIKFNGFNI 249
QY 142 ----RKLNLSNLVALG 153
Db 250 FFEIRYKEELLIVLG 265

Query Match 9.4%; Score 85; DB 2; Length 146;
Best Local Similarity 28.8%; Pred. No. 1.6;
Matches 36; Conservative 19; Mismatches 36; Indels 34; Gaps 7;
QY 1 MNKLLIFVLATFCVFSFAQANDSKNGAFG--MSAGEKLLVYETSKQDPI---VPF---- 51
Db 16 VTRIVTFILTSFVYLFPAHNDPGGFIIGGLMTASAFLLMYLAFDWDMSKKALPFNFTA 75
QY 52 -----LNLFLGFGIGSFAQGD-----ILGSLI---LGFDAVGIGLILAG 89
Db 76 LIAIGLLAIFT--GVSSMLAGDPFLTYQRYFQLPILGETELTALPFD-LGIYLVWIG 132
QY 90 AYLDI 94
Db 133 IALTI 137
RESULT 13

F83814
Nat/H+ antiporter BH1318 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
R:Accession: F83814
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: F83814
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-146 <STO>
A:Cross-references: GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BAB05037.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH1318

Query Match 9.4%; Score 85; DB 2; Length 146;
Best Local Similarity 28.8%; Pred. No. 1.6;
Matches 36; Conservative 19; Mismatches 36; Indels 34; Gaps 7;
QY 1 MNKLLIFVLATFCVSSFAANDSKNGAFG--MSAGEKLLVYETSKODPI---VPPF----- 51
Db 16 VTRIVTPIILSFYSYLFAGNDPGGFIFGGLMTASAFLLMLAFDMSMKKALPFFNFTA 75
QY 52 -----LLNLFGLFGIGSFAGQD-----ILGSLI---LGFDVAGIGLILAG 89
Db 76 LIAIGLLLAIFT--GVSSMLAGDPFLQYFRYFQLPILGETELTALPFD-LGIYLVVIG 132
QY 90 AYLDI 94
Db 133 IALTI 137

RESULT 14
C64142
conserved hypothetical protein HI0091 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 29-Sep-1999
C:Accession: C64142
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: C64142
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-378 <TIGR>
A:Cross-references: GB:U32694; GB:L42023; NID:g1573035; PIDN:AAC21769.1; PID:g1573042; T
C:Superfamily: yhaD protein

Query Match 9.4%; Score 84.5; DB 2; Length 378;
Best Local Similarity 28.3%; Pred. No. 4.6;
Matches 34; Conservative 19; Mismatches 48; Indels 19; Gaps 5;
QY 14 VFSSFAQANDSKNGAFGMSAGEKLLVYETSKODPIVPF-----LLNLFGLFGIGSFAGQ 67
Db 73 VKSEFGLSGDGKTAIEMAASGLHVPPEKRNPLLTSTVGTGELIKLALDLGVESFILG 132
QY 68 DILGGSILGFDAVIGLILAGAYLDIKALDGITKKAQFQWTGKGVMLAGVWTVMTVRL 127
Db 133 --IGGS---ATNDGGVGMLOA---LGMQCLDSQDKPIGF-----GGAELANIVKIDVQQL 179

RESULT 15
H70346
undecaprenyl-phosphate-alpha-N- acetylglucosaminyltransferase - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 17-Mar-2000

C:Accession: H70346
R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: H70346
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-357 <AQF>
A:Cross-references: GB:AE000694; NID:g2983162; PIDN:AAC06768.1; PID:g2983167; GB:AE0
C:Genetics:
A:Gene: rfe
C:Superfamily: lipophilic protein lim

Query Match 9.3%; Score 84; DB 2; Length 357;
Best Local Similarity 32.3%; Pred. No. 4.8;
Matches 32; Conservative 15; Mismatches 34; Indels 18; Gaps 4;
QY 14 VFSSFAQANDSK-----NGAFGMSAGEKLLV-----YETSKODPIVPFLLNLF-----LG 58
Db 139 LFTAFALAGVSHAFNIIIDGFNGLASGVAMLVFGAYAVSVFLHNDLFLVYLLNLLLIAATLG 198
QY 59 FGIGSFAGQDIL---GGSLILGFDAVIGLILAGAYLDI 94
Db 199 FFLMNYPFGLIFIGDGGAYFLGFCAGAILVKNYPDI 237

Search completed: March 14, 2003, 14:25:39
Job time : 15.4026 secs

THIS PAGE BLANK (USPTO)

197 PGLAFGAMASTIKYGV-----GSLVLQGLQILCYLTCILFFVLVLGTIAKENGFINI 249

QY 142 ----RKLNKSLINVALG 153

DB 250 FKFIYRIKEELLIVLG 265

RESULT 3

| ID | YHCL_BACSU | STANDARD; | PRT; | 463 AA. |
|----|--|-----------|------|---------|
| AC | P54596; | | | |
| DT | 01-OCT-1996 (Rel. 34, Created) | | | |
| DT | 01-OCT-1996 (Rel. 34, Last sequence update) | | | |
| DE | 15-JUN-2002 (Rel. 41, Last annotation update) | | | |
| DT | Hypothetical symporter yhcL | | | |
| GN | YHCL | | | |
| OS | Bacillus subtilis. | | | |
| OC | Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus. | | | |
| OX | NCBI_TaxID=1423; | | | |
| RN | [1] | | | |
| RN | SEQUENCE FROM N.A. | | | |
| RP | STRAIN=168; | | | |
| RC | MEDLINE=97124185; PubMed=8969498; | | | |
| RX | Noback M.A., Terpstra P., Holsappel S., Venema G., Bron S.; | | | |
| RA | "A 22 kb DNA sequence in the <i>cpbB</i> - <i>glpPKD</i> region at 75 degrees on the | | | |
| RA | Bacillus subtilis chromosome."; | | | |
| RA | Microbiology 142:3021-3026(1996). | | | |
| RL | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=168; | | | |
| RX | MEDLINE=98044033; PubMed=9384377; | | | |
| RA | Kunst F., Ogasawara N., Mosser I., Albertini A.M., Alloni G., | | | |
| RA | Azevedo V., Bertsier M.G., Bessieres P., Bolotin A., Borchert S., | | | |
| RA | Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S., | | | |
| RA | Brouillet S., Bruschii C.V., Caldwell B., Capuano V., Carter N.M., | | | |
| RA | Choi S.K., Codani J.K., Conerton I.F., Cummings N.J., Daniel R.A., | | | |
| RA | Denizot F., Devine K.M., Dusterhoft A., Enrich S.D., Emmerson P.T., | | | |
| RA | Entlan K.D., Errington J., Fabret C., Ferrari E., Foulger D., | | | |
| RA | Fritz C., Fujita Y., Fuma S., Galizzi A., Galleron N., | | | |
| RA | Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grand G., | | | |
| RA | Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A., | | | |
| RA | Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L., | | | |
| RA | Joris B., Karamata D., Kasahara Y., Klaerr-Bianchard M., Klein C., | | | |
| RA | Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M., | | | |
| RA | Kurita K., Lapidus A., Lardinot S., Lauber J., Lazarevic V., | | | |
| RA | Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C., | | | |
| RA | Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M., | | | |
| RA | Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., | | | |
| RA | Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M., | | | |
| RA | Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S., | | | |
| RA | Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadaie Y., | | | |
| RA | Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F., | | | |
| RA | Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B., | | | |
| RA | Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K., | | | |
| RA | Takeuchi M., Takakoshi A., Tanaka T., Terpstra P., Tognoni A., | | | |
| RA | Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A., | | | |
| RA | Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T., | | | |
| RA | Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., | | | |
| RA | Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A., | | | |
| RT | "The complete genome sequence of the Gram-positive bacterium <i>Bacillus</i> | | | |
| RT | <i>subtilis</i> ."; | | | |
| RT | Nature 390:249-256(1997). | | | |
| CC | -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable). | | | |
| CC | -!- SIMILARITY: BELONGS TO THE SODIUM:DICARBOXYLATE SYMPORTER FAMILY | | | |
| CC | (SDF, TC 2.A.23). | | | |
| CC | ---- | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL Outstation | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announcement/) | | | |

```
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X96983; CAA65696.1; -.
CC EMBL; Z99108; CAB12741.1; -.
CC Subtilist; BG11590; yhcL.
CC InterPro; IPR001991; Na/dico_symp.
CC Pfam; PF00375; SDF; 1.
CC PROSITE; PS00713; NA_DICARBOXYL_SYMP_1; FALSE_NEG.
CC PROSITE; PS00714; NA_DICARBOXYL_SYMP_2; FALSE_NEG.
CC KW Hypothetical protein; Transmembrane; Transport; Complete proteome.
CC FT TRANSMEM 3 23 POTENTIAL.
CC FT TRANSMEM 34 54 POTENTIAL.
CC FT TRANSMEM 73 93 POTENTIAL.
CC FT TRANSMEM 105 125 POTENTIAL.
CC FT TRANSMEM 184 204 POTENTIAL.
CC FT TRANSMEM 225 245 POTENTIAL.
CC FT TRANSMEM 262 282 POTENTIAL.
CC FT TRANSMEM 284 304 POTENTIAL.
CC FT TRANSMEM 338 358 POTENTIAL.
CC FT TRANSMEM 369 389 POTENTIAL.
CC FT TRANSMEM 394 414 POTENTIAL.
CC SEQUENCE 463 AA; 48982 MW; FA69EEAF5EC45F89 CRC64;

Query Match 9.5%; Score 85.5; DB 1; Length 463;
Best Local Similarity 24.3%; Pred. No. 2.7;
Matches 28; Conservative 25; Mismatches 43; Indels 19; Gaps 4;

Qy 27 GAFGMSAGEK-----LLVYETSKQDPIVP-FLNLFLFGIGSPFAGQDILGG-- 72
Db 334 GSFGLSGQNCAGTYPAMLAAMTAPTVGQNPFDVFITIAVVAISSFVGAVGGGAT 393
Qy 73 -SLIIGSDAVGIGLILAGAYLDIKALDGIITKKAQFQWTKGKVMLAGVVTAVTR 126
Db 394 FAALLVSLNMPALAGLLISIEPLIDMGRTALNV-----SGSMTSLGILTSKYK 444

RESULT 4
GRK_HAEIN
ID GRK_HAEIN STANDARD; PRT; 378 AA.
AC P44507;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glycinate kinase (EC 2.7.1.31).
GN GLXK OR HI0091.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OC NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RA MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
Rd.";
RL Science 269:496-512(1995).
CC -1- CATALYTIC ACTIVITY: ATP + (R)-glycerate -> ADP + 3-phospho-(R)-
CC glycerate.
CC -1- SIMILARITY: BELONGS TO THE GLYCERATE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U36839; AAB00548.1; -.
CC EMBL; U36839; AAB00548.1; -.

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32694; AAC21769.1; -.
CC TIGR; HI0091; -.
CC InterPro; IPR004381; Cons_hypoth45.
CC InterPro; IPR003747; Glycerate_kinase.
CC Pfam; PF02595; DUF168; 1.
CC TIGRFAMS; TIGR00045; Cons_hypoth45; 1.
CC Transferrase; Kinase; Complete proteome.
CC KW TRANSFERASE; Kinase; Complete proteome.
CC SEQUENCE 378 AA; 39758 MW; EBC57937BD8838A9 CRC64;

Query Match 9.4%; Score 84.5; DB 1; Length 378;
Best Local Similarity 28.3%; Pred. No. 2.7;
Matches 34; Conservative 19; Mismatches 48; Indels 19; Gaps 5;

Qy 14 VFSSFAQANDSKNGAFMSAGEKLLVYETSKQDPIVP-----LLNLFLFGIGSPFAGQ 67
Db 73 VKSFGLSGDGKTAIEMAAASGLHLVPPPEKRNPLITTSYGTGELIKLALDLGVESFIIG 132
Qy 68 DILGSLILGFDVAVGIGLILAGAYLDIKALDGIITKKAQFQWTKGKVMLAGVVTAVTR 127
Db 133 --IGGS---ATNDGGVGLQA---LGMQCLDSQDKPIGF-----GGAELANIVKIDVQOL 179

RESULT 5
FLIR_TREPA
ID FLIR_TREPA STANDARD; PRT; 265 AA.
AC P74932;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Flagellar biosynthetic protein flir.
GN FLIR OR TP0716.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OC NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RA MEDLINE=96105201; PubMed=8529894;
RA Hardham J.M., Frye J.G., Stamm L.V.;
RT "Identification and sequences of the Treponema pallidum flir", flir,
RT flir, flir, flir and flir genes";
RL Gene 166:57-64(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RA MEDLINE=9665876;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete";
RL Science 281:375-388(1998).
CC -1- FUNCTION: ROLE IN FLAGELLAR BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE FLIR/MOPE/SPAR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U36839; AAB00548.1; -.
CC EMBL; U36839; AAB00548.1; -.

```

DR EMBL; AE001244; AAC65682.1; -.
 DR TIGR; TP0716; -.
 DR InterPro; IPR002010; Bac_export_1.
 DR Pfam; PF01311; Bac_export_1; 1.
 DR PRINTS; PR00953; TYPE3IMPRROT.
 KW Flagella; Transmembrane; Complete proteome.
 FT TRANSMEM 14 34 POTENTIAL.
 FT TRANSMEM 44 64 POTENTIAL.
 FT TRANSMEM 73 93 POTENTIAL.
 FT TRANSMEM 132 152 POTENTIAL.
 FT TRANSMEM 182 202 POTENTIAL.
 FT TRANSMEM 229 249 POTENTIAL.
 SQ SEQUENCE 265 AA; 29313 MM; A9BFA16FE365C09 CRC64;

Query Match 9.1%; Score 82.5; DB 1; Length 265;
 Best Local Similarity 25.7%; Pred. No. 2.9;
 Matches 48; Conservative 27; Mismatches 63; Indels 49; Gaps 9;

OY 5 LIFVLATF---CVFSSPAQANDSKNGAFGMSAGEKLLVYETSKQ--DPIVFPFLN----- 54
 DB 80 LLGILTGFTTSVITFTPSAAGQFFSYQMFGTSE--MYDTFAQIENPLMGQFLNFVAML 136
 OY 55 -----LFLGFGIGSFAQGDIL-----GGSLLIGF-----DAVGIGLILAGA 90
 DB 137 VFLQIKGFQLLFGVLRSLRQVNCVFRLRKQKALLFFTKALSALFLHAMTIALPTMGA 196
 OY 91 YLDIKALDGTITKKAQFOWT-WKGGVLAGVVTMAVTRLTETIILPFTFANSYNRKLNLSN 149
 DB 197 LLLHVSMLLTAKAQMNLSEGLPTIVTVFV--LLSVILPY-----MINLFVS 245
 OY 150 VALGGFE 156
 DB 246 ILFGGFE 252

RESULT 6
 NUGM_RABIT STANDARD; PRT; 174 AA.
 AC 079438;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DE NADH-ubiquinone oxidoreductase chain 6 (EC 1.6.5.3).
 GN MTND6 OR ND6 OR NADH6.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98317530; PubMed=9653643;
 RA Gissi C., Gullberg A., Arnason U.;
 RT "The complete mitochondrial DNA sequence of the rabbit, *Oryctolagus cuniculus*".
 RL Genomics 50:161-169(1998).
 CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AJ001588; CA04858.1; -.
 DR InterPro; IPR001457; Oxidored_g3.
 DR Pfam; PF00499; oxidored_g3; 1.
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
 SQ SEQUENCE 174 AA; 18728 MM; 18740BB6661D09E1 CRC64;

Query Match 9.0%; Score 81.5; DB 1; Length 174;

Best Local Similarity 27.6%; Pred. No. 2.4;
 Matches 32; Conservative 18; Mismatches 35; Indels 31; Gaps 6;

OY 49 VPFLNLN-----FLGF-----GIGSPAQDILGSLILGFDVAVGILILAGAYL-D 93
 DB 4 VVLLSYNFMVGVFGFSKPSPIYGGGLIVSGGV-CGGIVLSFGSFLGLMFLIYLG 62
 OY 94 IKALDGTITKKA---FOWTWGKGYMLAG-----VVTMAVTRLTETIILPF 134
 DB 63 MLVVGFTTAMATEEYETWGSNMILGMFVLGVLMEVGLVVMVMSDGVETIVDF 118

RESULT 7
 PUC_RHOCA STANDARD; PRT; 461 AA.
 AC P23452;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Protein pucC.
 GN PUC.
 OS Rhodobacter capsulatus (Rhodospseudomonas capsulata).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 OC Rhodobacter.
 OX NCBI_TaxID=1061;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89359127; PubMed=2549005;
 RA Tichy H.V., Oberle B., Stiehle H., Schiltz E., Drews G.;
 RT "Genes downstream from pucB and pucA are essential for formation of
 RT the B800-850 complex of Rhodobacter capsulatus".
 RL J. Bacteriol. 171:4914-4922(1989).
 RN [2]
 RP REVISION, AND FUNCTION.
 RX MEDLINE=92007739; PubMed=1717257;
 RA Tichy H.V., Albien K.-U., Gad'On N., Drews G.;
 RT "Analysis of the Rhodobacter capsulatus puc operon: the pucC gene
 RT plays a central role in the regulation of LHII (B800-850 complex)
 RT expression".
 RL EMBO J. 10:2949-2955(1991).
 RN [3]
 RP TOPOLOGY.
 RX MEDLINE=96326322; PubMed=8759841;
 RA LeBlanc H.N., Beatty J.T.;
 RT "Topological analysis of the Rhodobacter capsulatus pucC protein and
 RT effects of C-terminal deletions on light-harvesting complex II".
 RL J. Bacteriol. 178:4801-4806(1996).
 CC -!- FUNCTION: PUC IS REQUIRED FOR HIGH-LEVEL TRANSCRIPTION OF THE
 CC PUC OPERON.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M28510; AAA26163.1; ALT_INIT.
 DR PIR; C33958; C33958.
 DR InterPro; IPR004896; PUC.
 DR Pfam; PF03209; PUC; 1.
 KW Antenna complex; Transmembrane.
 FT DOMAIN 1 36 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 37 56 PROBABLE.
 FT DOMAIN 57 62 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 63 83 PROBABLE.
 FT DOMAIN 84 109 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 110 129 PROBABLE.
 FT DOMAIN 130 142 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 143 163 PROBABLE.
 FT DOMAIN 164 182 CYTOPLASMIC (PROBABLE).


```
CC CC (SDF, TC 2.A.23).
CC -!- CAUTION: There are two genes for DctA in NGR234; one on the sym
CC plasmid, the other on the chromosome.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S38912; AAB22400.1; ALT_INIT.
CC DR EMBL; S38912; AAB22401.1; -.
CC DR EMBL; Z68203; CAA92421.1; ALT_INIT.
CC DR EMBL; AE000101; AAB91894.1; ALT_INIT.
CC DR InterPro; IPR001991; Na/dico_symp.
CC DR Pfam; PF00375; SDF; 1.
CC DR PROSITE; PS00713; NA_DICARBOXYL_SYMP_1; 1.
CC DR PROSITE; PS00714; NA_DICARBOXYL_SYMP_2; 1.
CC KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
CC KW Plasmid.
CC FT TRANSMEM 27 45 POTENTIAL.
CC FT TRANSMEM 60 82 POTENTIAL.
CC FT TRANSMEM 95 117 POTENTIAL.
CC FT TRANSMEM 168 185 POTENTIAL.
CC FT TRANSMEM 205 227 POTENTIAL.
CC FT TRANSMEM 237 259 POTENTIAL.
CC FT TRANSMEM 348 370 POTENTIAL.
CC FT TRANSMEM 374 393 POTENTIAL.
CC FT CONFLICT 101 101 T -> A (IN REF. 1).
CC FT CONFLICT 337 337 A -> G (IN REF. 1).
CC FT CONFLICT 437 497 PVETMGLSQRSQSTVELGKQLVFGATNSADRTLAGRPGGR
CC DSRRAPDHSQAQVFGPLSL -> AGGDHGVGVKPAE (IN
CC REF. 1).
CC SQ SEQUENCE 497 AA; 52128 MW; 49E960D8845B2824 CRC64;
Query Match 8.9%; Score 80.5; DB 1; Length 497;
Best Local Similarity 26.9%; Pred. No. 7.9;
Matches 52; Conservative 26; Mismatches 70; Indels 45; Gaps 11;
QY 3 KLLIFVL--ATFCVFSSFAQAANDSKNGAFGMSA-----GEKLIVY-ETSKQDPFVPFL 53
DB 94 KAMIFYLTFTALLVLGVVAVVQPCA-GMHIDPASLDAKATATYA-EKAHEQSVTGFLM 152
QY 54 NLFLGFIGSGFAOGDIL-----GGSL-ILGFDAVGIGLILAGAYLDIKALDGITKK 103
DB 153 NIITPTLVGAFAEGDIQLQVLFISVFGISLAIVGKKAEEVVDFLHALTLPIFRVLAIKM 212
QY 104 A-----AQFWTWGK-GVMLAGVTVNAVTRLEIILPPTFA-----NYSN---- 141
DB 213 AAPTAGFAFMATIGTKYGV--ASTIANIAMLIGTFYLSFLFVFMVLGAVARYNGFSIVAL 270
QY 142 -RKLNLSNLVALG 153
DB 271 IRVIKEELLIVLG 283
RESULT 10
ID YA96_MYCPN STANDARD; PRT; 264 AA.
AC PF75556;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MPN096 (R02_orf264).
GN MPN096 OR MP058.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
NCBI_TaxId=2104;
RN [1]
RC STRAIN=ATCC 29342 / W129.
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:22:51 ; Search time 24.47 Seconds
(without alignments)
1507.249 Million cell updates/sec

Title: US-09-508-487-19
Perfect score: 902
Sequence: 1 MNKLIFVLATFCVFSFAQ.....DVAMGSSALGFELSFKSKY 179

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 902 | 100.0 | 179 | 16 | O51065 borrelia bu |
| 2 | 769.5 | 85.3 | 178 | 2 | Q9X309 |
| 3 | 737 | 81.7 | 177 | 2 | Q9X3V0 |
| 4 | 291 | 32.3 | 161 | 16 | O50896 |
| 5 | 279.5 | 31.0 | 170 | 16 | O50885 |
| 6 | 247.5 | 27.4 | 161 | 2 | Q9S011 |
| 7 | 235 | 26.1 | 190 | 16 | O50696 |
| 8 | 155.5 | 17.2 | 95 | 2 | O31325 |
| 9 | 94.5 | 10.5 | 427 | 2 | Q93RE6 |
| 10 | 91 | 10.1 | 616 | 16 | Q8UAV6 |
| 11 | 87.5 | 9.7 | 198 | 16 | Q8KK8 |
| 12 | 87.5 | 9.7 | 289 | 16 | Q9XK6 |
| 13 | 87.5 | 9.7 | 442 | 17 | Q976Y7 |
| 14 | 87.5 | 9.7 | 462 | 16 | Q92MI7 |
| 15 | 86.5 | 9.6 | 1874 | 2 | O33763 |
| 16 | 86 | 9.5 | 214 | 16 | Q55742 |

| | | | | | |
|----|------|-----|------|----|--------|
| 17 | 86 | 9.5 | 274 | 16 | O8XQB4 |
| 18 | 86 | 9.5 | 303 | 2 | Q93PN3 |
| 19 | 86 | 9.5 | 307 | 2 | Q9XB19 |
| 20 | 85 | 9.4 | 146 | 16 | O44659 |
| 21 | 85 | 9.4 | 995 | 5 | Q9N3H7 |
| 22 | 85 | 9.4 | 355 | 10 | Q9FIS2 |
| 23 | 84 | 9.3 | 357 | 16 | O66804 |
| 24 | 83.5 | 9.3 | 470 | 5 | Q19932 |
| 25 | 83.5 | 9.3 | 538 | 17 | Q9HH12 |
| 26 | 83 | 9.2 | 420 | 16 | Q9CP19 |
| 27 | 82.5 | 9.1 | 438 | 17 | Q96XR2 |
| 28 | 82.5 | 9.1 | 439 | 2 | Q9RBQ9 |
| 29 | 82 | 9.1 | 153 | 17 | O8TYN4 |
| 30 | 82 | 9.1 | 553 | 11 | Q9Z306 |
| 31 | 82 | 9.1 | 1021 | 17 | Q8TWM0 |
| 32 | 81.5 | 9.0 | 296 | 17 | Q9HL22 |
| 33 | 81.5 | 9.0 | 336 | 16 | P73160 |
| 34 | 81 | 9.0 | 144 | 2 | Q9RG24 |
| 35 | 81 | 9.0 | 279 | 16 | O87858 |
| 36 | 81 | 9.0 | 427 | 2 | Q8VS09 |
| 37 | 81 | 9.0 | 457 | 2 | Q9R637 |
| 38 | 81 | 9.0 | 508 | 16 | Q9PNA1 |
| 39 | 81 | 9.0 | 553 | 11 | Q9RI41 |
| 40 | 81 | 9.0 | 695 | 17 | Q97A18 |
| 41 | 80.5 | 8.9 | 455 | 16 | Q9ZS78 |
| 42 | 80.5 | 8.9 | 650 | 16 | Q97RE5 |
| 43 | 80.5 | 8.9 | 666 | 16 | O9A523 |
| 44 | 80.5 | 8.9 | 709 | 16 | P73100 |
| 45 | 80.5 | 8.9 | 1914 | 2 | O33762 |

ALIGNMENTS

RESULT 1

| | | | |
|--------|--|------|---------|
| O51065 | PRELIMINARY; | PRT; | 179 AA. |
| AC | O51065; | | |
| DT | 01-JUN-1998 (TrEMBLrel. 06, Created) | | |
| DT | 01-JUN-1998 (TrEMBLrel. 06, Last sequence update) | | |
| DT | 01-MAR-2002 (TrEMBLrel. 20, Last annotation update) | | |
| DE | Hypothetical protein BB0034. | | |
| GN | BB0034 OR P13. | | |
| OS | Borrelia burgdorferi (Lyme disease spirochete). | | |
| OC | Bacteria; Spirochaetales; Spirochaetaceae; Borrelia. | | |
| OX | NCBI_taxid=139; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=ATCC 35210 / B31; | | |
| EX | MEDLINE=98065943; PubMed=9403685; | | |
| RA | Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A., | | |
| RA | Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M., | | |
| RA | Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D., | | |
| RA | Peterson J., Kervatage A.R., Quackenbush J., Salzberg S., Hanson M., | | |
| RA | van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J., | | |
| RA | Utterback T., Matthey L., McDonald L., Artiach P., Bowman C., | | |
| RA | Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B., | | |
| RA | Smith H.O., Venter J.C.; | | |
| RT | "Genomic sequence of a Lyme disease spirochete, Borrelia | | |
| RL | burgdorferi." | | |
| RN | Nature 390:580-586(1997). | | |
| RP | [2] | | |
| RC | SEQUENCE FROM N.A. | | |
| RC | STRAIN=ATCC 35210 / B31; | | |
| RT | Noppa L., Ostberg Y., Bergstrom S.; | | |
| RT | "p13, a novel integral membrane protein of Lyme disease Borrelia." | | |
| RL | Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases. | | |
| DR | EMBL; AF001117; AAC66426.1; - | | |
| DR | EMBL; AF085739; AAD28360.1; - | | |
| DR | TIGR; BB0034; - | | |
| DR | Hypothetical protein; Complete proteome. | | |
| SQ | SEQUENCE 179 AA; 19104 MW; FD6056E8E24E5D5D CRC64; | | |

```
Query Match 100.0%; Score 902; DB 16; Length 179;
Best Local Similarity 100.0%; Pred. No. 2.8e-72;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKLLIFVLATFCVFSFAQANDSKNGAFGMSAGEKLLVYETSKQDPIVPFLLNLFGLFG 60
DB 1 MNKLLIFVLATFCVFSFAQANDSKNGAFGMSAGEKLLVYETSKQDPIVPFLLNLFGLFG 60

QY 61 IGSAQGDILGSLILGFDVAGIGLILAGAYLDIKALDGIITKKAQFQWTGKGMVLGAV 120
DB 61 IGSAQGDILGSLILGFDVAGIGLILAGAYLDIKALDGIITKKAQFQWTGKGMVLGAV 120

QY 121 TMAVTRTEIILPFTFANSYNRKLKNSLNVALGFGFSPFDVAMQSSALGFLSFKKSY 179
DB 121 TMAVTRTEIILPFTFANSYNRKLKNSLNVALGFGFSPFDVAMQSSALGFLSFKKSY 179

RESULT 2
Q9X3U9 PRELIMINARY; PRT; 178 AA.
AC Q9X3U9;
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=29518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ACAI;
RX MEDLINE=21189251; PubMed=11292755;
RA Noppa L., Ostberg Y., Lavrinovicha M., Bergstrom S.;
RT "p13, an integral membrane protein of Borrelia burgdorferi, is C-
terminally processed and contains surface-exposed domains.";
RL Infect. Immun. 69:3323-3334(2001).
DR EMBL: AF085740; AAD28361.1; -.
SQ SEQUENCE 178 AA; 19194 MW; C1C032EB5320C86 CRC64;

Query Match 85.3%; Score 769.5; DB 2; Length 178;
Best Local Similarity 84.9%; Pred. No. 1.5e-60;
Matches 152; Conservative 10; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNKLLIFVLATFCVFSFAQANDSKNGAFGMSAGEKLLVYETSKQDPIVPFLLNLFGLFG 60
DB 1 MNKLLIFVLATFCVFSFAQANDSKS-AFNLGAGEKLLAYETSKQDPIVPFLLNLFGLFG 59

QY 61 IGSAQGDILGSLILGFDVAGIGLILAGAYLDIKALDGIITKKAQFQWTGKGMVLGAV 120
DB 60 IGSAQGDILGSLILGFDVAGIGLILAGAYLDIKALDKNAPKAFKWTGKGMVLGAV 119

QY 121 TMAVTRTEIILPFTFANSYNRKLKNSLNVALGFGFSPFDVAMQSSALGFLSFKKSY 179
DB 120 TMAVTRTEIILPFTFANSYNRKLKNSLNATAGFGFSPFDINMGQASALGFLSFKKSY 178

RESULT 3
Q9X3V0 PRELIMINARY; PRT; 177 AA.
AC Q9X3V0;
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=29519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ip90;
RX MEDLINE=21189251; PubMed=11292755;
```

```
RA Noppa L., Ostberg Y., Lavrinovicha M., Bergstrom S.;
RT "p13, an integral membrane protein of Borrelia burgdorferi, is C-
terminally processed and contains surface-exposed domains.";
RL Infect. Immun. 69:3323-3334(2001).
DR EMBL: AF085741; AAD28362.1; -.
SQ SEQUENCE 177 AA; 19308 MW; 429928CC426E1111 CRC64;

Query Match 81.7%; Score 737; DB 2; Length 177;
Best Local Similarity 81.0%; Pred. No. 1.1e-57;
Matches 145; Conservative 14; Mismatches 18; Indels 2; Gaps 2;

QY 1 MNKLLIFVLATFCVFSFAQANDSKNGAFGMSAGEKLLVYETSKQDPIVPFLLNLFGLFG 60
DB 1 MNKFLIFILVIFCAFSSFAQ-DDSKS-TFNLGAGEFLVYETNKKDSLVPFLLNLFGLFG 58

QY 61 IGSAQGDILGSLILGFDVAGIGLILAGAYLDIKALDGIITKKAQFQWTGKGMVLGAV 120
DB 59 IGSAQGDILGSLILGFDVAGIGLILAGAYLDIKDFDNNAKRADFKWTGKGMVLGAV 118

QY 121 TMAVTRTEIILPFTFANSYNRKLKNSLNVALGFGFSPFDVAMQSSALGFLSFKKSY 179
DB 119 TMAVTRTEIILPFTFANNYNRKLKNSLNALGFGFSPFDINMGQASALGFLSFKKSY 177

RESULT 4
O50896 PRELIMINARY; PRT; 161 AA.
AC O50896;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein BBA01.
GN BBA01.
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp54
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
Peterson J., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
Utterback T., Wattney L., McDonald L., Artiach P., Bowman C.,
Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
burgdorferi.";
RL Nature 390:580-585(1997).
DR EMBL: AE000790; AAC66226.1; -.
DR TIGR: BBA01; -.
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 161 AA; 17887 MW; D93BDB326FE2DA30 CRC64;
```

```
Query Match 32.3%; Score 291; DB 16; Length 161;
Best Local Similarity 39.5%; Pred. No. 3e-18;
Matches 75; Conservative 23; Mismatches 42; Indels 50; Gaps 8;

QY 4 LLIFVLATFCVFSFAQANDSKNGAFGMSAGEKLLVYETSKQDPIVPFLLNLFGLFG 63
DB 8 LLIFGL-TQIFATKDTQNRKEGI-----ESFNKYDREKKNIPGFLNLFGLFG 60

QY 64 FAQGDILGSLILGFDVAGIGLILAGAYLDIKALDGIITKKAQFQWTGKGV----- 114
DB 61 FVGQDYIGGSGVLGN-----LLGAIL-----WGTLILNHRQTOL 96

QY 115 ---MLAGV-VTMAVTRLT-EIILPFTFANSYNRKLKNSLNVALGFGFSPFDVAMQSSAL 169
DB 97 TGVILGCVASWLTSTVSLIIPFTFANNHNENLKRSLAELAGFEPNLDGIN----- 151
```



```
QY 170 GFELSFKKS 179
DB 152 GFQLSFKKS 161

RESULT 5
OS0885 PRELIMINARY; PRT; 170 AA.
AC OS0885;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein BBI31.
GN BBI31.
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp28-4.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kervatage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uterback T., Watthey L., McDonald L., Artlich P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi."
RL Nature 390:580-586(1997).
DR EMBL; AE000789; AAC66194.1; -
DR TIGR; BBI31; -
RW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 170 AA; 18227 MW; 0E71B633EA80FB22 CRC64;

Query Match 31.0%; Score 279.5; DB 16; Length 170;
Best Local Similarity 36.9%; Pred. No. 3.4e-17;
Matches 66; Conservative 29; Mismatches 69; Indels 15; Gaps 4;

QY 1 MNKLLIFVATFCVFSFAQAANDSKNGAFMSAGEKLLVYETSKODPIVPFLLNLFGLG 60
DB 7 MKKLTLLIFSLPIQIFATQDKLEK---SVGSITIMKYKSEKATILAPLLNLFGLG 63
QY 61 IGSAQDGLGSLILGFDVAGIGLILAGAYLDIKALDITKKAQFQWTKGKVMLAGV 120
DB 64 IGSPVQGDYIGGGAVLGSLGLICIAAGNLG---HTDDETRA-----TTGHIITTTG 116
QY 121 TMAYVTRLTEILPFTFANSYNRKLNSLVALGSGFSPFSDVAMGSSALGFELSKSY 179
DB 117 TIIASHASLIIPFTFANKHNANLKKRLGIDIAGFEPNFDIGIS-----GFQLSFKKRY 170

RESULT 6
Q9S011 PRELIMINARY; PRT; 161 AA.
AC Q9S011;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Conserved hypothetical protein.
GN BQ06.
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp56.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Lathigra R.,
RA White O., Dodson R., Hickey E.K., Gwinn M., Peterson J.,
RA Palmer N., Haft D., Rosa P., Stevenson B.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi."
RL Nature 390:580-586(1997).
DR EMBL; AE000784; AAC65989.1; -
DR TIGR; BQ06; -
RW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 161 AA; 21813 MW; A80FB8BCAD4B13EB CRC64;

Query Match 26.1%; Score 235; DB 16; Length 190;
Best Local Similarity 38.9%; Pred. No. 3.3e-13;
Matches 56; Conservative 25; Mismatches 51; Indels 12; Gaps 4;

QY 37 LLVYETSKQ-DPIVPFLLNLFGLGIGSFAQDILGSLILGFDVAGIGLILAGAYLDIK 95
DB 58 LLQYKEQGLIDLWIPVMLNLFDPFGVSGVDYIGGCTGLGNLGLTGTG---VI 113
QY 96 ALDGIITKKAQFQWTKGKVMAGVYVMAVTRILTEILPFTFANSYNRKLNSLVALG 155
DB 114 QMKNLKKEPAS--ISSMILLSGMLTFGSSYLSIYLPVLPFDYRYKLMNRIDELAG 171
QY 156 EPSFDVAMGQSSALGFELSKSY 179
DB 172 EPNLDIGMN-----GFQLSFKKSY 190

"A bacterial genome in flux: The twelve linear and nine circular
extrachromosomal DNAs in an infectious isolate of the Lyme disease
spirochete Borrelia burgdorferi.",
Mol. Microbiol. 0:0-0(1999).
DR EMBL; AE001584; AAF07707.1; -
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 161 AA; 17016 MW; 14D4906CC8107CD2 CRC64;

Query Match 27.4%; Score 247.5; DB 2; Length 161;
Best Local Similarity 40.9%; Pred. No. 2.2e-14;
Matches 63; Conservative 25; Mismatches 51; Indels 15; Gaps 6;

QY 18 FAQANDSKNGAFMSAGEKLLVYETSKODPIVPFLLNLFGLGIGSFAQDILGSLILG 77
DB 4 FAQEKLEK---GVGDIATVMKYSEKATILAPLLNLFGLGIGSFAQDILGSLILG 59
QY 78 FDVAGIGLILAGAYL--DIKALDGIITKKAQFQWTKGKVMLAGV--VTMAVTRLTEILP 133
DB 60 SQVLGGILIMAG-YMTGDI---GFVTESTATVITGG---VLSGIGGLTAAASYITGIT 112
QY 134 FTFANSYNRKLNSLVALGSGFSPFSDVAMGSS 167
DB 113 FKFNRYNADUKKRLGIALAGLEPNFDIGNGDS 146

RESULT 7
OS0696 PRELIMINARY; PRT; 190 AA.
ID OS0696
AC OS0696;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein BHH41.
GN BHH41.
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp28-3.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A. / B31;
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kervatage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uterback T., Watthey L., McDonald L., Artlich P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi."
RL Nature 390:580-586(1997).
DR EMBL; AE000784; AAC65989.1; -
DR TIGR; BHH41; -
RW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 190 AA; 21813 MW; A80FB8BCAD4B13EB CRC64;
```

```

99 GI--TKKAQFQWTKGKVMGLAGVVTMAVTRLTETIILPFTFANSYNRKLNLSNALGGFE 156
187 GYLWLKXHLIAFGWD--IRASLVLDKTRFLSLSLPTTF-----NFLAWAGLGF- 234

QY 157 PSFDVAMGQSSALG 170
DB 235 -AYHAIMGQAGVGQ 247

RESULT 10
Q8UAV6
ID Q8UAV6 PRELIMINARY; PRT; 616 AA.
AC Q8UAV6;
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DE ABC transporter, membrane spanning protein.
GN ATU3261 OR AGR_L_3113.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmli K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
DR EMBL; AE009255; AAL44077.1; -
DR EMBL; AE008358; AAK90129.1; -
SQ Complete proteome.
KW SEQUENCE 616 AA; 63402 MW; A6005D4347ABE060 CRC64;

Query Match 10.1%; Score 91; DB 16; Length 616;
Best Local Similarity 22.2%; Pred. No. 7.4;
Matches 42; Conservative 33; Mismatches 76; Indels 38; Gaps 8;

QY 9 LATFCVFSFAQANDSKNGAFGMSAGEKLLVYE-----TSKQDPIVP----- 50
DB 301 LUTSVLFSG-ASGSSVANAFGASTQPELVKHGYPGAPAGAIATSVLDNVIPPSIAF 359
QY 51 FLINLFLGFGIGSFAQGDILGSLILGFDVAVGIGLILAGAYLDIKALDGIK- KAAQW- 108
DB 360 LILATATNLVSGSLLVGGFFAGGLM---AVCLGVAI---HLSVRSVDTLPRTAGRW 412
QY 109 -----TWGKGMGLAGVVTMAVTRLTETIILPFTFANSYNRKLNLSNALGGFE 161
DB 413 SIAIAIPAGLGVVVVVGIRIGIVTTTAA---ALAAALYLLGLFGYRLGVGRIFATFRQ 469
QY 162 AMGQSSALG 170
: |:::|:

QY 1 MKKLFVLVATCFVSSFAQANDSKNGAFGMSAGEKLLVYETSKQDPIVPFLNLFLGFG 60
DB 1 MKKIFTLILFGLTIEFATKDTQNRKIEGIESFNK---YDREKNPIGPFLLNLFLPG 57
QY 61 IGSFAQGDILGSLILGFDVAVG 82
DB 58 IGSFVGQDYIGGSGVLGFLNLLG 79

RESULT 9
Q93RE6
ID Q93RE6 PRELIMINARY; PRT; 427 AA.
AC Q93RE6;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Adhesin.
GN VAG.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RA Akeda Y., Honda T.;
RT "Vibrio parahaemolyticus adhesin gene."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047560; BAB59008.1; -
DR InterPro; IPR002528; Mate.
DR Pfam; PF01554; URF0013; 2.
DR TIGRfams; TIGR00797; mate; 1.
SQ SEQUENCE 427 AA; 45909 MW; 0B608B9DC70E3C71 CRC64;

Query Match 10.5%; Score 94.5; DB 2; Length 427;
Best Local Similarity 25.8%; Pred. No. 2.4;
Matches 50; Conservative 25; Mismatches 72; Indels 47; Gaps 11;

QY 5 LIFVLATCFVSSFAQ-----ANDSKNGAFGMSAGEKLLVYETSKQDPIV----- 49
DB 73 MVFAALVCLFVFFQPIGLTDSOE---VIELGSSYLVISSASFAVACVASMAGLR 129
QY 50 ----PFLNLFLGFGIGS--FQGDILGSLILGFDVAVG-----TGLILAGAYLDIKALD 98
DB 130 AMHQPGLSFTFFSGIGLSNVLNWLFIQGH--LGFPALGITGAATATVISGA-IEVGCLF 186
```

```
Db 470 SAGEAAIG 478
RESULT 11
Q98KK8 PRELIMINARY; PRT; 198 AA.
AC Q98KK8;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein m11430.
GN M11430.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RL DNA Res. 7:331-338(2000).
DR EMBL; AP002997; BAB48806.1; -.
DR InterPro; IPR001123; Lyse.
DR Pfam; PF01810; Lyse; 1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 198 AA; 21126 MW; 67E14425F02E7DAB CRC64;

Query Match 9.7%; Score 87.5; DB 16; Length 198;
Best Local Similarity 30.0%; Pred. No. 4;
Matches 45; Conservative 17; Mismatches 51; Indels 37; Gaps 9;

Qy 49 VPFLNLFLGFGIGSFAOGDIUG-GSLILGFDVAVGIGL-ILAGAVL-----DIKALDGIT 101
Db 40 VPHMLGISGFLVLLAVG--LGAGVLTAPFALTALKIAGGAYLLYLAKWIAMRSRLT 97
Qy 102 KK-----AAFQWTWKGVMGLAGVVTMAVTRLTE-----IILPFFANSY 140
Db 98 GKGETKARPMRFIDAAAFQWNPRAWNA-ITAMAVYANAEPFLSVALISTAFIVNLP 156
Qy 141 NRKLKNSLUNVALGGF--EPS-----FDVANG 164
Db 157 SVSVWAGFTALRGFLSDPVLKWFNIANG 186

RESULT 12
Q9KXK6 PRELIMINARY; PRT; 289 AA.
AC Q9KXK6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative integral membrane protein.
GN SC02314 OR SC053.05.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL357591; CAB93434.1; -.
SQ SEQUENCE 289 AA; 29123 MW; 89AD68D9D805BE10 CRC64;

Query Match 9.7%; Score 87.5; DB 16; Length 289;
Best Local Similarity 32.6%; Pred. No. 6.2;
Matches 31; Conservative 12; Mismatches 37; Indels 15; Gaps 4;

Qy 73 SLILGFDVAVGIG---LILAGAYLDIKALDGITKKAQFQWTWKGVMGLAGVVTMAVTRLTE 129
Db 5 SLTRG-DGVVIGAAVLLLIASFLDIYSIDGAPDSADIPSLWGSVGLVGLAGVIGAL 63

Qy 130 IILPFTFANSYNRKLKNSLUNVA---LGGFEPSPFDV 161
Db 64 VVV-----NRAMPQPKVAGLDLGGFGIAFTV 90

RESULT 13
Q976Y7 PRELIMINARY; PRT; 442 AA.
AC Q976Y7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Putative transporter.
GN ST0054.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kawabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
DR EMBL; AP000981; BAB65009.1; -.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 442 AA; 47602 MW; C7932E5ABF2F03A CRC64;

Query Match 9.7%; Score 87.5; DB 17; Length 442;
Best Local Similarity 22.8%; Pred. No. 10;
Matches 47; Conservative 29; Mismatches 75; Indels 55; Gaps 10;

Qy 4 LLIFVLATFCVF-SSFAQANDSKNGA--FGMSAGEKL-----LVYETSKQD-- 46
Db 49 VFLATAVYATFLTSLM--RPVGLYFGNSISDKLGRKRLIYIGLIVITASLQAL 105
Qy 47 -----PIVPELNLFLGFGIGSFAQGD-----ILGGSILILGFDVAG 82
Db 106 PTYAVVGFAPILILVRLAEFFVGGVTAGSHITGPESVPERHGWVG--IGFSAAG 162
Qy 83 IGLIILAGAYLDIKALDGITKKAQFQWTWKGVM-LAGVVTMAVTRLTEIILP----FTFA 137
Db 163 AAYLTASAWFYMTAL---LFGSGSYLVMGWRVYMFEGGLPLAVLGVSVLYVPESDAWQKA 219
Qy 138 NSYNRKLKNSLUNVALGGFEPSPFDVAM 163
Db 220 KDRGKTIKSPVRELFKFRFTFGIAL 245
```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:19:56 ; Search time 30.8258 Seconds
(without alignments) 765.116 Million cell updates/sec

Title: US-09-508-487-23
 Perfect score: 908
 Sequence: 1 MNKELIFILVIFCAFSSFAQ.....DINMGQASALGFLSFKSY 177

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seas, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*

| | |
|-----|---|
| 1: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.* |
| 2: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.* |
| 3: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.* |
| 4: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.* |
| 5: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.* |
| 6: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.* |
| 7: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.* |
| 8: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.* |
| 9: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.* |
| 10: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.* |
| 11: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.* |
| 12: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.* |
| 13: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.* |
| 14: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.* |
| 15: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.* |
| 16: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.* |
| 17: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.* |
| 18: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.* |
| 19: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.* |
| 20: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.* |
| 21: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.* |
| 22: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.* |
| 23: | /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 908 | 100.0 | 177 | 20 | Borrelia garinii I |
| 2 | 785.5 | 86.6 | 178 | 20 | Borrelia afzelii A |
| 3 | 737 | 81.2 | 179 | 20 | B. burgdorferi ant |
| 4 | 737 | 81.2 | 179 | 20 | Borrelia burgdorfe |
| 5 | 667.5 | 73.5 | 158 | 20 | B. burgdorferi ant |
| 6 | 170.5 | 18.8 | 95 | 20 | p11 protein sequen |
| 7 | 111 | 12.2 | 25 | 20 | Borrelia burgdorfe |
| 8 | 88.5 | 9.7 | 451 | 22 | Novel mar regulate |
| 9 | 88.5 | 9.7 | 451 | 22 | Escherichia coli p |
| 10 | 87 | 9.6 | 413 | 22 | Botulinum toxin hea |

| | | | | | | |
|----|------|-----|------|----|----------|---------------------|
| 11 | 87 | 9.6 | 848 | 22 | AAB04082 | Botulism toxin hea |
| 12 | 87 | 9.6 | 852 | 23 | ABG69077 | Botulinum neurotox |
| 13 | 87 | 9.6 | 858 | 19 | AAW56017 | Recombinant botuli |
| 14 | 87 | 9.6 | 1070 | 21 | AAV93308 | A manganese supero |
| 15 | 87 | 9.6 | 1095 | 21 | AAV93311 | A manganese supero |
| 16 | 87 | 9.6 | 1169 | 19 | AAV56017 | Recombinant botuli |
| 17 | 87 | 9.6 | 1291 | 19 | AAW68392 | Clostridium botuli |
| 18 | 82 | 9.0 | 1456 | 22 | ABG07111 | Novel human diagno |
| 19 | 82 | 9.0 | 1456 | 22 | ABG24514 | Novel human diagno |
| 20 | 81.5 | 9.0 | 206 | 22 | AAU35869 | Helicobacter pylori |
| 21 | 80 | 8.8 | 222 | 20 | AAV34545 | Porphorymonas ging |
| 22 | 80 | 8.8 | 227 | 20 | AAV34411 | Porphorymonas ging |
| 23 | 79.5 | 8.8 | 333 | 23 | ABP27264 | Streptococcus poly |
| 24 | 79.5 | 8.8 | 726 | 22 | AAU55177 | Propionibacterium |
| 25 | 79 | 8.7 | 479 | 20 | AAV19965 | B. burgdorferi ant |
| 26 | 79 | 8.7 | 504 | 22 | AAV19964 | B. burgdorferi ant |
| 27 | 78 | 8.6 | 398 | 22 | AAV78998 | C. glutamicum SRT |
| 28 | 78 | 8.6 | 544 | 22 | AAV90173 | C. glutamicum prote |
| 29 | 78 | 8.6 | 544 | 22 | AAV78997 | C. glutamicum SRT |
| 30 | 78 | 8.6 | 544 | 23 | AAE19943 | C. glutamicum gl |
| 31 | 77 | 8.5 | 243 | 23 | ABP38116 | Corynebacterium gl |
| 32 | 76.5 | 8.4 | 345 | 22 | AAV78781 | Staphylococcus epi |
| 33 | 76.5 | 8.4 | 417 | 22 | ABG93096 | Corynebacterium gl |
| 34 | 76 | 8.4 | 237 | 22 | ABG27217 | Novel human diagno |
| 35 | 76 | 8.4 | 283 | 18 | AAW55463 | Novel human diagno |
| 36 | 76 | 8.4 | 293 | 18 | AAW55244 | H. pylori ORF 05ae |
| 37 | 76 | 8.4 | 456 | 22 | ABG58293 | H. pylori ORF 05ae |
| 38 | 76 | 8.4 | 781 | 22 | ABG26754 | Drosophila melanog |
| 39 | 75.5 | 8.3 | 392 | 22 | AAU67004 | Novel human diagno |
| 40 | 75.5 | 8.3 | 904 | 21 | AAV51119 | Propionibacterium |
| 41 | 75.5 | 8.3 | 935 | 15 | AAV35552 | E. coli 0-157 vero |
| 42 | 75.5 | 8.3 | 935 | 15 | AAW55550 | EHEC gene prod |
| 43 | 75 | 8.3 | 215 | 19 | AAW99808 | Amino acid sequenc |
| 44 | 75 | 8.3 | 474 | 23 | AAE16782 | H. pylori GPHO 132 |
| 45 | 75 | 8.3 | 998 | 18 | AAV33621 | Human transporter |
| 46 | 75 | 8.3 | 45 | 23 | AAV33621 | Elmeria tenella 45 |

ALIGNMENTS

| | |
|----------|--|
| RESULT 1 | |
| AA04280 | |
| IID | AA04280 standard; Protein; 177 AA. |
| XX | |
| AC | AA04280; |
| DT | 17-JUN-1999 (first entry) |
| XX | |
| DE | Borrelia garinii IP30 protein. |
| XX | |
| KW | Borrelia; P13 antigen; spirochaete; |
| XX | |
| OS | Borrelia garinii. |
| XX | |
| PN | WO9912960-A2. |
| XX | |
| PD | 18-MAR-1999. |
| XX | |
| PF | 04-SEP-1998; 98WO-IB01424. |
| XX | |
| PR | 16-SEP-1997; 97US-0059036. |
| XX | |
| PR | 10-SEP-1997; 97DK-0001041. |
| XX | |
| PP | (SYMB-) SYMBICOM AB. |
| XX | |
| PI | Bergstroem S; |
| XX | |
| DR | WPI: 1999-215027/18. |
| XX | |
| DR | N-PSDB; AAX30100. |
| XX | |
| PT | Nucleic acid from Borrelia burgdorferi |
| XX | |
| PPT | protein P13 |
| XX | |

PT caused by *Borrelia*, particularly Lyme disease

PS Claim 12; Page 72; 275pp; English.

This sequence represents a *Borrelia burgdorferi* (Bb) protein of the invention, which is suitable for use in a vaccine. The Bb polypeptides can be used in vaccines for eliciting protective antibodies to members of the *Borrelia* genus, particularly for the use against Lyme disease in humans and animals. They can be used for preventing or attenuating infection caused by a member of the *Borrelia* genus. The products can also be used for detection of members of the *Borrelia* genus.

| Sequence | 179 AA; |
|----------|---------|
| SQ | |

| | | | | |
|-----------------------|------------------|--------------------|--------|-------------|
| Query Match | 81.2%; | Score 737; | DB 20; | Length 179; |
| Best Local Similarity | 81.0%; | Pred. No. 2.7e-73; | | |
| Matches 145; | Conservative 14; | Mismatches 18; | Indels | |

Oy

1 MNKFLIFILVIFCAFSSFAQ-DDSKS-TFNLGAGEKFVYETNKKDSLVFPFNLFLGFG 58
||| ||| : | | | | | : | : | | | | |
||| ||| : | | | | | : | : | | | | |

Dd

1 MNKLIIIVLATFCVFSSFAQANDSKNGAFGMSAGEKLLVYTSTKODIPVPFNLFLGFG 60

59 IGSFAQGDILGGSLLGFDAVGIGLILTGAYLDIKDFEDNNAKKADFKWTWGKMMLAGVV 11

Db 61 IGSFAQGDILGSLILGFDPAVGIGLILAGAYLDIKALDITKKAAFQWTWKGVMLAGVV 120

Qy 119 TMAVTRITEIVLPTTFANNYNRKLNLSNIALGGFEPFSDINMQASALGFLGFLSPKKS 177

Db 121 TMAVTRLTEIILPFTFANSYNRKLSNLVALGGFEPFDVANGQSSALGFELSPKSY 179

RESULT 4
AAY04278
ID AAY04278 standard: Protein: 179 AA.

XX
AC
XX

DT 17-JUN-1999 (first entry)

XX
DE
Borrellia burgdorferi B31 protein[illegible]

KW Borrellia;

XX
05 Borrelia burgdorferiXX
XXXXXX
XXXXXXXXXX

PN WO9912960-A2.

XX
19-MAR-1990

FD-302 (Rev. 1-25-60)

04-SEP-1998;

16-00000-1007-0340 00500000

| | | |
|----|--------------|---------------|
| PR | 16-SEP-1997; | 97US-0059036. |
| PR | 10-SEP-1997: | 97DK-0001041 |

XX
XX

PA (SYMB-) SYMBICOM AB.

XX
PT
Perastrom C.XX
perg. 3,

DR WPI; 1999-215027/

DR N-PSDB; AAX30098.
XX
PT Nucleic acid from *Borrelia burgdorferi* encoding virulence associated
PM

protein PI3

PS
CC The present sequence represents a Borrelia burgdorferi B313 protein.
CC The present invention describes an isolated nucleic acid (I) that: (i)
CC encodes a polypeptide fragment (II) immunologically reactive with
CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of
CC Borrelia burgdorferi B313, but not reactive with most proteins from
CC other spirochaetes; and/or (ii) hybridizes under stringent conditions
CC to specific nucleic acid sequences but not to genomic DNA from most

Claim 31; Page 108-109; 118pp; English.

CC humans and animals. They can be used for preventing or attenuating an
CC infection caused by a member of the *Borrelia* genus. The products can also
CC be used for detection of members of the *Borrelia* genus.

SQ Sequence 158 AA;
Query Match 73.5%; Score 667.5; DB 20; Length 158;
Best Local Similarity 82.3%; Pred. No. 1.1e-65;
Matches 130; Conservative 13; Mismatches 14; Indels 1; Gaps 1;
QY 21 DSKS-TFNLGAGEKFLVYETNKKDSLVPFLNLFGLFGIGSFAQDDILGSLILGFDAY 79
Db 1 NDSKNGAFGSGAGEKLLVYETKQDPVFPFLNLFGLFGIGSFAQDDILGSLILGFDAY 60
QY 80 GTGLITGAYLDIKDFDNNAKADKFTWKGKMMLAGVVTMAVTRTEVLPTFANNYN 139
Db 61 GTGLILAGAYLDIKALDGTITKAAFTWKGKMMLAGVVTMAVTRTEVLPTFANNYN 120
140 RLKLSNLALGGFEPSPDINMGQASALGFLSFKKSY 177
Db 121 RLKLSNLVALGGFEPSPDVAMQSSALGFLSFKKSY 158

RESULT 6
AAY04282
ID AAY04282 standard; Protein; 95 AA.
XX
AC AAY04282;
XX
DT 17-JUN-1999 (first entry)
XX
DE p11 protein sequence.
XX
KW *Borrelia*; P13 antigen; spirochaete; vaccine; infection; diagnosis.
XX
OS Unidentified.
XX
PN WO9912960-A2.
XX
PD 18-MAR-1999.
XX
PF 04-SEP-1998; 98WO-IB01424.
XX
PR 16-SEP-1997; 97US-0059036.
PR 10-SEP-1997; 97DK-0001041.
XX (SYMB-) SYMBICOM AB.

PI Bergstroem S;
XX
DR WPI; 1999-215027/18.
DR N-PSDB; AAX30115.
XX
PT Nucleic acid from *Borrelia burgdorferi* encoding virulence associated
PT protein P13
XX
PS Example 8; Page 118; 118pp; English.
XX
CC The present invention describes an isolated nucleic acid (I) that: (i)
CC encodes a polypeptide fragment (II) immunologically reactive with
CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of
CC *Borrelia burgdorferi* B313, but not reactive with most proteins from
CC other spirochaetes; and/or (ii) hybridizes under stringent conditions
CC to specific nucleic acid sequences but not to genomic DNA from most
CC other spirochaetes. (I), (II) and transformed cells are useful in
CC vaccines to protect against infection by *B. burgdorferi* sensu lato.
CC (I), (II) and antibodies raised against (II) are used to diagnose such
CC infections, in standard immunoassays or amplification/hybridization
CC tests. (I) are also used to produce recombinant (II). The 13 kD protein
CC is involved in virulence and is highly conserved within *B. burgdorferi*
CC sensu lato, but is absent from other *Borrelia* species (e.g. those
CC responsible for relapsing fever or avian borreliosis). It should provide
CC a more specific and sensitive antibody response, and diagnosis,

CC compared with use of whole bacteria as antigen. The present sequence
CC represents a p11 protein sequence from a best hit search.

SQ Sequence 95 AA;
Query Match 18.8%; Score 170.5; DB 20; Length 95;
Best Local Similarity 45.0%; Pred. No. 4.6e-11;
Matches 36; Conservative 11; Mismatches 32; Indels 1; Gaps 1;
QY 1 MKKFLIFILVIFCAFFSSFAQDDSKSTFNLGAGEKFLVYETNKKDSLVPFLNLFGLFGIG 60
Db 1 MKKIFTLILIFGLTIEIFATKDTQNIKGI-ESFNKYDKERKNPIGPFLLNLFPLFGIG 59
QY 61 SFAQGDILGSLILGFDAY 80
Db 60 SFVQGDYIGGSGVLGNLLG 79

RESULT 7
AAY04281
ID AAY04281 standard; peptide; 25 AA.
XX
AC AAY04281;
XX
DT 17-JUN-1999 (first entry)
XX
DE *Borrelia burgdorferi* B313 N-terminal peptide.
XX
KW *Borrelia*; P13 antigen; spirochaete; vaccine; infection; diagnosis.
XX
OS *Borrelia burgdorferi*.
XX
PN WO9912960-A2.
XX
PD 18-MAR-1999.
XX
PF 04-SEP-1998; 98WO-IB01424.
XX
PR 16-SEP-1997; 97US-0059036.
PR 10-SEP-1997; 97DK-0001041.
XX (SYMB-) SYMBICOM AB.
XX
PI Bergstroem S;
XX
DR WPI; 1999-215027/18.
XX
PT Nucleic acid from *Borrelia burgdorferi* encoding virulence associated
PT protein P13
XX
PS Example 6; Page 49; 118pp; English.

XX The present sequence represents a *Borrelia burgdorferi* B313 N-terminal
CC peptide. The present invention describes an isolated nucleic acid (I)
CC that: (i) encodes a polypeptide fragment (II) immunologically reactive
CC with rabbit polyclonal antibody raised against a 13 kDa polypeptide of
CC *Borrelia burgdorferi* B313, but not reactive with most proteins from
CC other spirochaetes; and/or (ii) hybridizes under stringent conditions
CC to specific nucleic acid sequences but not to genomic DNA from most
CC other spirochaetes. (I), (II) and transformed cells are useful in
CC vaccines to protect against infection by *B. burgdorferi* sensu lato.
CC (I), (II) and antibodies raised against (II) are used to diagnose such
CC infections, in standard immunoassays or amplification/hybridization
CC tests. (I) are also used to produce recombinant (II). The 13 kD protein
CC is involved in virulence and is highly conserved within *B. burgdorferi*
CC sensu lato, but is absent from other *Borrelia* species (e.g. those
CC responsible for relapsing fever or avian borreliosis). It should provide
CC a more specific and sensitive antibody response, and diagnosis,

SQ Sequence 25 AA;
Query Match 12.2%; Score 111; DB 20; Length 25;

THIS PAGE BLANK (USPTO)

Qy 121 AATAGTTCACAAAGCCCTTTTAAATTCATTAATATGCTACCAATAATACCAGTTTAATA 180
Db 121 AATAGTTCACAAAGCCCTTTTAAATTCATTAATATGCTACCAATAATACCAGTTTAATA 180
Qy 181 AAGGGGTTTTATGAATAAGTTTTTAAATTTTATTTTGGTAAATCTTTTGGCTTTTCTA 240
Db 181 AAGGGGTTTTATGAATAAGTTTTTAAATTTTATTTTGGTAAATCTTTTGGCTTTTCTA 240
Qy 241 GTTTTGTCTCAAGATGATTCATAAAGCACTTTTAAATCTGGGAGCGGAGAAAATTTTGG 300
Db 241 GTTTTGTCTCAAGATGATTCATAAAGCACTTTTAAATCTGGGAGCGGAGAAAATTTTGG 300
Qy 301 TTTATGAACCTAATAAAGAAATCTCTGTGACCAATTTTATTTGAACCTTTTTTATAGGGT 360
Db 301 TTTATGAACCTAATAAAGAAATCTCTGTGACCAATTTTATTTGAACCTTTTTTATAGGGT 360
Qy 361 TCGGGATAGGTTCTTTTGTCTCAAGGAGATATCCTTTGGAGGTTCTCTTATTTCTGGATTG 420
Db 361 TCGGGATAGGTTCTTTTGTCTCAAGGAGATATCCTTTGGAGGTTCTCTTATTTCTGGATTG 420
Qy 421 ATGCGGTTGGTATAGGGTTAATCTTACAGGAGCTTTATTTGGACATCAAGGATTTTGATA 480
Db 421 ATGCGGTTGGTATAGGGTTAATCTTACAGGAGCTTTATTTGGACATCAAGGATTTTGATA 480
Qy 481 ATAATGCTAAAAAGCTGATTTTAAAGTGGACTTGGGTTAAGGAAATGATGTTGGCAGGTG 540
Db 481 ATAATGCTAAAAAGCTGATTTTAAAGTGGACTTGGGTTAAGGAAATGATGTTGGCAGGTG 540
Qy 541 TGGTTACTATGGCTGACAAAGATTTGACAGAAATTTGTTCTTCCATTTACATTTGCTAATA 600
Db 541 TGGTTACTATGGCTGACAAAGATTTGACAGAAATTTGTTCTTCCATTTACATTTGCTAATA 600
Qy 601 ATTATACAGAAAGCTGAAAAATAGTCTTAATATAGCTTGGAGGATTTTCAGCCTAGTT 660
Db 601 ATTATACAGAAAGCTGAAAAATAGTCTTAATATAGCTTGGAGGATTTTCAGCCTAGTT 660
Qy 661 TTGATATTAAACATGGGCCAAGCTAGTCTCTTGGTTTGGACTGCTCTTTCAAAAAAGCT 720
Db 661 TTGATATTAAACATGGGCCAAGCTAGTCTCTTGGTTTGGACTGCTCTTTCAAAAAAGCT 720
Qy 721 ATTAATTTTATTTATCTAGAAAATGGGTG 749
Db 721 ATTAATTTTATTTATCTAGAAAATGGGTG 749

RESULT 3

US-09-508-487-22

; Sequence 22, Application US/09508487

; GENERAL INFORMATION:

; APPLICANT: BERGSTROM, Sven

; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA

; FILE REFERENCE: 454312-3150

; CURRENT APPLICATION NUMBER: US/09/508,487

; PRIOR FILING DATE: 2000-03-10

; PRIOR APPLICATION NUMBER: 60/059,036

; PRIOR FILING DATE: 1997-09-10

; PRIOR APPLICATION NUMBER: 1041/97

; PRIOR FILING DATE: 1997-09-10

; PRIOR APPLICATION NUMBER: PCT/IB98/01424

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 22

; LENGTH: 749

; TYPE: DNA

; ORGANISM: Borrelia garinii

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (192)..(725)

US-09-508-487-22

Query Match

Best Local Similarity 100.0%; Score 749; DB 19; Length 749;

Pred. No. 4.4e-122;

Matches 749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAGAATTTTCAACAAATAAGATATTGTTAAAAGAAATTTGAAATGCTAATTTTATGGTTA 60
Db 1 TAGAATTTTCAACAAATAAGATATTGTTAAAAGAAATTTGAAATGCTAATTTTATGGTTA 60
Qy 61 AATCAAGAAGCTCTATTGGTAAAGCGAATTTTCGAGTAACAATTTGAAAAAGTTAAATTTA 120
Db 61 AATCAAGAAGCTCTATTGGTAAAGCGAATTTTCGAGTAACAATTTGAAAAAGTTAAATTTA 120
Qy 121 AATAGTTCACAAAGCCCTTTTAAATTCATTAATATGCTACCAATAATACCAGTTTAATA 180
Db 121 AATAGTTCACAAAGCCCTTTTAAATTCATTAATATGCTACCAATAATACCAGTTTAATA 180
Qy 181 AAGGGGTTTTATGAATAAGTTTTTAAATTTTATTTTGGTAAATCTTTTGGCTTTTCTA 240
Db 181 AAGGGGTTTTATGAATAAGTTTTTAAATTTTATTTTGGTAAATCTTTTGGCTTTTCTA 240
Qy 241 GTTTTGTCTCAAGATGATTCATAAAGCACTTTTAAATCTGGGAGCGGAGAAAATTTTGG 300
Db 241 GTTTTGTCTCAAGATGATTCATAAAGCACTTTTAAATCTGGGAGCGGAGAAAATTTTGG 300
Qy 301 TTTATGAACCTAATAAAGAAATCTCTGTGACCAATTTTATTTGAACCTTTTTTATAGGGT 360
Db 301 TTTATGAACCTAATAAAGAAATCTCTGTGACCAATTTTATTTGAACCTTTTTTATAGGGT 360
Qy 361 TCGGGATAGGTTCTTTTGTCTCAAGGAGATATCCTTTGGAGGTTCTCTTATTTCTGGATTG 420
Db 361 TCGGGATAGGTTCTTTTGTCTCAAGGAGATATCCTTTGGAGGTTCTCTTATTTCTGGATTG 420
Qy 421 ATGCGGTTGGTATAGGGTTAATCTTACAGGAGCTTTATTTGGACATCAAGGATTTTGATA 480
Db 421 ATGCGGTTGGTATAGGGTTAATCTTACAGGAGCTTTATTTGGACATCAAGGATTTTGATA 480
Qy 481 ATAATGCTAAAAAGCTGATTTTAAAGTGGACTTGGGTTAAGGAAATGATGTTGGCAGGTG 540
Db 481 ATAATGCTAAAAAGCTGATTTTAAAGTGGACTTGGGTTAAGGAAATGATGTTGGCAGGTG 540
Qy 541 TGGTTACTATGGCTGACAAAGATTTGACAGAAATTTGTTCTTCCATTTACATTTGCTAATA 600
Db 541 TGGTTACTATGGCTGACAAAGATTTGACAGAAATTTGTTCTTCCATTTACATTTGCTAATA 600
Qy 601 ATTATACAGAAAGCTGAAAAATAGTCTTAATATAGCTTGGAGGATTTTCAGCCTAGTT 660
Db 601 ATTATACAGAAAGCTGAAAAATAGTCTTAATATAGCTTGGAGGATTTTCAGCCTAGTT 660
Qy 661 TTGATATTAAACATGGGCCAAGCTAGTCTCTTGGTTTGGACTGCTCTTTCAAAAAAGCT 720
Db 661 TTGATATTAAACATGGGCCAAGCTAGTCTCTTGGTTTGGACTGCTCTTTCAAAAAAGCT 720
Qy 721 ATTAATTTTATTTATCTAGAAAATGGGTG 749
Db 721 ATTAATTTTATTTATCTAGAAAATGGGTG 749

RESULT 4

US-09-153-447-20

; Sequence 20, Application US/09153447

; GENERAL INFORMATION:

; APPLICANT: BERGSTROM, Sven

; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA

; FILE REFERENCE: 454312-3130.1

; CURRENT APPLICATION NUMBER: US/09/153,447

; CURRENT FILING DATE: 1998-09-15

; EARLIER APPLICATION NUMBER: 60/059,036

; EARLIER FILING DATE: 1997-09-10

; EARLIER APPLICATION NUMBER: 1041/97

; EARLIER FILING DATE: 1997-09-10

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: PatentIn ver. 2.1

; SEQ ID NO 20

; LENGTH: 862

; TYPE: DNA

; ORGANISM: Borrelia afzelii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (219)..(755)
US-09-153-447-20

Query Match 81.0%; Score 607; DB 15; Length 862;
Best Local Similarity 90.1%; Pred. No. 4.3e-97;
Matches 673; Conservative 0; Mismatches 70; Indels 4; Gaps 2;

QY 1 TAGAATTTTCAACAATAAAGATATTTGTTAAAGAATTTGAAATGCTAAATTTTATGTTA 60
DB 27 TAGAATTTTTCGACAAATAAAGACATTTATTAAGAATTTGAAATGCTAAATTTTATGTTA 86
QY 61 AATCAAGAAGCTCTATTTGGTAAGCGAATTTTCGAGTAACAATTTGAAAAAGTTAAATTTA 120
DB 87 AATCAAGAAGCTCTATTTGGTAAGCGAATTTTCGAGTAACAATTTGAAAAAGTTAAATTTA 146
QY 121 AATAGTTCCAAAAGCGCTTTTAAATTTTCAATTAATGCTACCATTAATACCAG-TTTAAT 179
DB 147 AATAGTTTAAAGCGCTTTTAAATTTTCAATTAATGCTACCATTAATACCAGTTTAAAT 206
QY 180 AAGAGGTTTATGTAATAGTTTAAATTTTCAATTAATGCTACCATTAATACCAG-TTTAAT 239
DB 207 AAGAGGTTTATGTAATAGTTTAAATTTTCAATTAATGCTACCATTAATACCAGTTTAAAT 266
QY 240 AGTTTTCGCTCAA---GATGATTTCTAAAAGCGCTTTTAAATTTGAGCGGAGAGAAATTT 296
DB 267 AGTTTTCGCTCAAAGCTGATGATTTCTAAAAGCGCTTTTAAATTTGAGCGGAGAGAAATTT 326
QY 297 TTGGTTTATGAACTAATAGAAGATTCCTTTGTAACCATTTTATGTAACCTTTTAAATTTA 356
DB 327 TTAGCTTATGAACTAATAGAAGATTCCTTTGTAACCATTTTATGTAACCTTTTAAATTTA 386
QY 357 GGGTTTCGGATAGGTTCTTTCGCTCAAGAGATATCTTTGAGGTTCTCTTATTTCTTGA 416
DB 387 GGGTTTGGAAATAGGTTCTTTCGCTCAAGAGATATCTTTGAGGTTCTCTTATTTCTTGA 446
QY 417 TTGTGTCGGTTAGGTTCTTTCGCTCAAGAGATATCTTTGAGGTTCTCTTATTTCTTGA 476
DB 447 TTGTGTCGGTTAGGTTCTTTCGCTCAAGAGATATCTTTGAGGTTCTCTTATTTCTTGA 506
QY 477 GATAATATGCTAAGAGCTGATTTTTCGCTCAAGAGATATCTTTGAGGTTCTCTTATTTCTTGA 536
DB 507 GATAATATGCTAAGAGCTGATTTTTCGCTCAAGAGATATCTTTGAGGTTCTCTTATTTCTTGA 566
QY 537 GGTGTCGGTTAGGTTCTTTCGCTCAAGAGATATCTTTGAGGTTCTCTTATTTCTTGA 596
DB 567 GGTGTCGGTTAGGTTCTTTCGCTCAAGAGATATCTTTGAGGTTCTCTTATTTCTTGA 626
QY 597 AATAATATGCTAAGAGCTGATTTTTCGCTCAAGAGATATCTTTGAGGTTCTCTTATTTCTTGA 656
DB 627 AATAATATGCTAAGAGCTGATTTTTCGCTCAAGAGATATCTTTGAGGTTCTCTTATTTCTTGA 686
QY 657 AGTTTTCGATTAACATGGCCAGCTAGTCTCTTTGGTTTTCGAGTCTCTTCAAAAAA 716
DB 687 AGTTTTCGATTAACATGGCCAGCTAGTCTCTTTGGTTTTCGAGTCTCTTCAAAAAA 746
QY 717 AGCTATTAAATTTTATTTATCTAGAAAA 743
DB 747 AGTATTAAATTTTATTTATTTATTTAA 773

RESULT 5

US-09-156-447-20

; Sequence 20, Application US/09156447

; GENERAL INFORMATION:

; APPLICANT: BERGSTROM, Sven

; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA

; FILE REFERENCE: 454312-3130.1

; CURRENT APPLICATION NUMBER: US/09/156,447

; CURRENT FILING DATE: 1998-09-15

; PRIOR APPLICATION NUMBER: 60/059,036

; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 1041/97
; PRIOR FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; TYPE: DNA
; LENGTH: 862
; ORGANISM: Borrelia afzelii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (219)..(755)
US-09-156-447-20

Query Match 81.0%; Score 607; DB 15; Length 862;
Best Local Similarity 90.1%; Pred. No. 4.3e-97;
Matches 673; Conservative 0; Mismatches 70; Indels 4; Gaps 2;

QY 1 TAGAATTTTCAACAATAAAGATATTTGTTAAAGAATTTGAAATGCTAAATTTTATGTTA 60
DB 27 TAGAATTTTTCGACAAATAAAGACATTTATTAAGAATTTGAAATGCTAAATTTTATGTTA 86
QY 61 AATCAAGAAGCTCTATTTGGTAAGCGAATTTTCGAGTAACAATTTGAAAAAGTTAAATTTA 120
DB 87 AATCAAGAAGCTCTATTTGGTAAGCGAATTTTCGAGTAACAATTTGAAAAAGTTAAATTTA 146
QY 121 AATAGTTCCAAAAGCGCTTTTAAATTTTCAATTAATGCTACCATTAATACCAG-TTTAAT 179
DB 147 AATAGTTTAAAGCGCTTTTAAATTTTCAATTAATGCTACCATTAATACCAGTTTAAAT 206
QY 180 AAGAGGTTTATGTAATAGTTTAAATTTTCAATTAATGCTACCATTTTATGTAACCTTTTCT 239
DB 207 AAGAGGTTTATGTAATAGTTTAAATTTTCAATTAATGCTACCATTTTATGTAACCTTTTCT 266
QY 240 AGTTTTCGCTCAA---GATGATTTCTAAAAGCGCTTTTAAATTTGAGCGGAGAGAAATTT 296
DB 267 AGTTTTCGCTCAAAGCTGATGATTTCTAAAAGCGCTTTTAAATTTGAGCGGAGAGAAATTT 326
QY 297 TTGGTTTATGAACTAATAGAAGATTCCTTTGTAACCATTTTATGTAACCTTTTAAATTTA 356
DB 327 TTAGCTTATGAACTAATAGAAGATTCCTTTGTAACCATTTTATGTAACCTTTTAAATTTA 386
QY 357 GGTTCGGATAGGTTCTTTCGCTCAAGAGATATCTTTGAGGTTCTCTTATTTCTTGA 416
DB 387 GGTTCGGATAGGTTCTTTCGCTCAAGAGATATCTTTGAGGTTCTCTTATTTCTTGA 446
QY 417 TTGTGTCGGTTAGGTTCTTTCGCTCAAGAGATATCTTTGAGGTTCTCTTATTTCTTGA 476
DB 447 TTGTGTCGGTTAGGTTCTTTCGCTCAAGAGATATCTTTGAGGTTCTCTTATTTCTTGA 506
QY 477 GATAATATGCTAAGAGCTGATTTTTCGCTCAAGAGATATCTTTGAGGTTCTCTTATTTCTTGA 536
DB 507 GATAATATGCTAAGAGCTGATTTTTCGCTCAAGAGATATCTTTGAGGTTCTCTTATTTCTTGA 566
QY 537 GGTGTCGGTTAGGTTCTTTCGCTCAAGAGATATCTTTGAGGTTCTCTTATTTCTTGA 596
DB 567 GGTGTCGGTTAGGTTCTTTCGCTCAAGAGATATCTTTGAGGTTCTCTTATTTCTTGA 626
QY 597 AATAATATGCTAAGAGCTGATTTTTCGCTCAAGAGATATCTTTGAGGTTCTCTTATTTCTTGA 656
DB 627 AATAATATGCTAAGAGCTGATTTTTCGCTCAAGAGATATCTTTGAGGTTCTCTTATTTCTTGA 686
QY 657 AGTTTTCGATTAACATGGCCAGCTAGTCTCTTTGGTTTTCGAGTCTCTTCAAAAAA 716
DB 687 AGTTTTCGATTAACATGGCCAGCTAGTCTCTTTGGTTTTCGAGTCTCTTCAAAAAA 746
QY 717 AGCTATTAAATTTTATTTATCTAGAAAA 743
DB 747 AGTATTAAATTTTATTTATTTATTTAA 773

RESULT 6

US-09-508-487-20

| | | | |
|----|-----|---|-----|
| Db | 541 | GACAGATTACGAAATTAATCTTCATTTACATTTGCTAATAGTTATAATAGGAAGCT | 600 |
| Qy | 617 | GAATAATAGTCTTAATATACGCTTGGGAGGATTTGAGCCCTAGTATTTGATATTAACATGGG | 676 |
| Db | 601 | AAAAAATAGCCTTAATGTAGCTTTAGGAGGATTTGAACCTAGTATTTGATGCTTGCATATGGG | 660 |
| Qy | 677 | CCAAGCTAGTGCCTCTTGTTTGGTCTTGGACTGCTCTTCAAAAAGCTATTAAATTTATTTATC | 736 |
| Db | 661 | CCAATCAGTGCCTCTGGGTTTGAACCTGCTTTCAAAAAGCTATTAAATTTATTTATT | 720 |
| Qy | 737 | TAGAAATGGGTG | 749 |
| Db | 721 | ACAAAATGGGTG | 733 |

```

RESULT 8
US-09-156-447-18
; Sequence 18, Application US/09156447
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3130.1
; CURRENT APPLICATION NUMBER: US/09/156,447
; CURRENT FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/059,036
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 1041/97
; PRIOR FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Borrelia burgdorferi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (170)..(709)
US-09-156-447-18

```

| | | | | |
|-----------------------|--------------|--------------------|----------------|-------------|
| Query Match | 74.5% | Score 558.2; | DB 15; | Length 759; |
| Best Local Similarity | 87.0%; | Pred. No. 1.6e-88; | | |
| Matches 638; | Conservative | 0; | Mismatches 88; | Indels 7; |
| Gaps 2; | | | | |

| | | | |
|----|-----|---|--------|
| Qy | 24 | ATTGTTAAAAGAATGAAATGCTAAATTTATGGTTTAAATCAAGAAGCTCTATTGGTAA | 83 |
| Db | 1 | ATTGTTAAAAGAATGAAATTTGATAATTTATGGTCAAATCAAGAAGCTCTATTGGGA | 60 |
| Qy | 84 | CGAATTTTCGAGTAAACAATTTGAAAAAGTTAAATTTAAATAGTTCCAAAAGCCTTTT | 143 |
| Db | 61 | CGAATTTCAAGCAATAATTTGAAAAAGTTAAATTTAAATACTTTAAAAACCTTTT | 120 |
| Qy | 144 | AAATTCATTAATATGCTACCATAAATACAG-TTTAATAAAGGGTTTTTATGAATAAGTT | 202 |
| Db | 121 | AAATTTTCATTAATATGCTACCATAGTACCAGTTTTAATAAGGGTTTTTATGAATAA | 180 |
| Qy | 203 | TTTAAATTTTATTTTGGTAATCTTTTGTGCTTTTCTAGTTTGTGTCAGAATGATTC | -- 260 |
| Db | 181 | TTTAAATTTTGTTTTGGCAACCTTTTGTGTTTTTCTAGCTTTGTCTCAAGCTAATG | 240 |
| Qy | 261 | ----AAAAGCACHTTTAATCTGGGACGGGAGAAAAATTTTGGTTTATGAACCTAA | 316 |
| Db | 241 | TAAAAATGGTGCCTTTTGGATAGTCTCGAGAAAAACCTTTTGGTTTATGAACCTA | 300 |
| Qy | 317 | GAAGATCTCTGTGTAACCATTTTATTGAACCTTTTTTAGGGTTCGGATAGGTCCTT | 376 |
| Db | 301 | GCAGATPCTTATTTGTAACCATTTTATTGAACCTTTTTTAGGGTTTGGAAATAGG | 360 |
| Qy | 377 | TGCTCAAGGAGATATCCTTGGAGGTCTCTATTCTTGGATTGTATGCGGTGGTATAG | 436 |
| Db | 361 | TGCTCAAGGAGATTTCTTGGAGGTCTCTATTCTTGGATTGTATGCGGTGGTATAG | 420 |
| Qy | 437 | GTTTAATACTTACAGGAGCTATTTTGGACATCAAGGATTTTGATAATAATGCTTAA | 496 |

[illegible]

RESULT 9
US-09-508-487-18
Sequence 18, Application US/09508487
GENERAL INFORMATION:
APPLICANT: BERGSTROM, Sven
TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
FILE REFERENCE: 454312-3150
CURRENT APPLICATION NUMBER: US/09/508,487
CURRENT FILING DATE: 2000-03-10
PRIOR APPLICATION NUMBER: 60/059,036
PRIOR FILING DATE: 1997-09-10
PRIOR APPLICATION NUMBER: 1041/97
PRIOR FILING DATE: 1997-09-10
PRIOR APPLICATION NUMBER: PCT/IB98/01424
PRIOR FILING DATE: 1998-09-04
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 18
LENGTH: 759
TYPE: DNA
ORGANISM: Borrelia burgdorferi
FEATURE:
NAME/KEY: CDS
LOCATION: (170)..(709)
US-09-508-487-18

Query Match 74.5%; Score 558.2; DB 19; Length 759;
Best Local Similarity 87.0%; Pred. No. 1.6e-88;
Matches 638; Conservative 0; Mismatches 88; Indels 7; Gaps 2;

| | | | |
|----|-----|--|-----|
| Qy | 24 | ATTGTTAAAAAGAAATTGAAATTGCTAAATTTTATGTTTAAATCAAGAAGCTCTATTGGTAAAG | 83 |
| Db | 1 | ATTGTTAAAAAGAAATTGAAATTGATAATTTTATGTCACAAATCAAGAAGCTCTATTGGGAAG | 60 |
| Qy | 84 | CGAATTCGAGTAGTACAAATTTGAAAAAGATTAAATTTAAATAGTTCAAAAAGCCTTTTTTA | 143 |
| Db | 61 | CGAATTTTCAAGCAATAAATTTGAAAAAGATTAAATTTAAATAACTTTTAAAAACCTTTTTTA | 120 |
| Qy | 144 | AATTTCTAATTAATGCTTACCATAATACCAAG-TTTTAATAAAGGGGTTTTTATGAATAAGTT | 202 |
| Db | 121 | AATTTCTAATTAATGCTTACCATAGTACCAGTTTAAATAAAGGGGTTTTTATGAATAAACT | 180 |
| Qy | 203 | TTTAAATTTTTTATTTTGGTAACTTTTGTGCTTTTTTCTAGTTTTCGTCACAGATGATTC-T | 260 |
| Db | 181 | TTTAAATTTTTTGTTTTGGCAACCTTTTGTGTTTTTCTAGCTTTTCGTCACAGCTAATGATTC | 240 |
| Qy | 261 | -----AAAAGCAGCTTTTAACTCGGGAGCGGGAGAAAAATTTTGGTTTATGAAACTAATAA | 316 |
| Db | 241 | TAAAAATGGTGGTTTTGGGATGAGTCTGCAGAAAAAATTTTGGTTTTTATGAACAATGACAA | 300 |

PCT-US98-12764-63

Query Match 13.18; Score 98.4; DB 1; Length 1761;
Best Local Similarity 49.78; Pred. No. 1.4e-07;
Matches 336; Conservative 0; Mismatches 331; Indels 9; Gaps 3;

1 TAGAATTTTCAACAAATAAAGATATTGTTAAAGAAATTTGAAATTTGCTAATTTTATGGTTA 60
Qy 1031 TGGCATACAAAAAATAGATAGCATAAATTTCAAAATTTAAAGAGATTTAAAGAAAATTT 972
Db 61 AATCAAGAGCTCTATTGGTAAGCGAATTTTCGAGTAACAATTTGAAAAAGTTAAATTTA 120
Qy 971 ACATTTAAATCTGCTTAGAACAGCAAGAAATTTCAACAATATCTTTTAAAGAAAGATTCA 912
Db 121 AATAGTTCCAAAGCCTTTTAAATTTCAATTAATGCTACCAATAATACCAAGTTAATA 180
Qy 911 ACTCCATAAATAAATCTAAAAATTTGCTTAACATTTGTTTATCAAAACTTTAATTTATTAG 852
Db 181 AAGG---GGTTTATGAATAAGTTTAAATTTTAAATTTTATTTTGGTAATCTTTTGTGCTTTT 237
Qy 851 GAGGTAATTAATTAATGAAGAAATTTTCACATTAATAATTTTATTTTGTAGTTTAAACAATGC 792
Db 238 CTAGTTTTCGCTCAAGATGATCTTAAAGCACTTTTAACTCGGAGCGGGAGAAAAATTTT 297
Qy 791 AAATCTTTGCACAAGATAAACTTTGAAAAAGGTGTTGGAGT-GGAGATATCGCGACTGTA 733
Db 298 TGGTTTATGAACATAAAGAAAGATCTCTGTGTACCAATTTTATTTGAACCTTTTATTAG 357
Qy 732 TGAATAATGAAGCCGAAAAAGCAACCACTACTAGCACCACTTACTTTTGAATATCTTTTGT 673
Db 358 GGTTCGGGATAGTTCTTTTGTCTCAAGGAGATATCTCTGGAGGTCTCTTATTTCTTGGAT 417
Qy 672 CTTTAGGATAGGATCTTTTGTTCAGGAGATTTATTTGGTGTGTGCTACTTGGCT 613
Db 418 TTGATCGCGGTGGTATAGGTTAACTTTACAGGAGCTTTATTTGGACATCAAGGATTTTG 477
Qy 612 CTCAGTGTCTGGAAG-----AATCTTATTAATGCTGGACACATGATCGTAGGAGGTA 558
Db 478 ATAATAATGCTAAAAAGCTGATTTTAACTGGACITGGGGTAAGGAAATGATGTTGGCAG 537
Qy 557 TTGGAGCGGTTCACAGAAAGCAGCAAGTAAACCGGAGGAGTATTTATCAGGAATAG 498
Db 538 GTGTGTTTACTATGCTGTGCAAGATTTGACAGAAATTTTCTTCCATTTCATTTGCTTA 597
Qy 497 GAGGCTCACAATTTAGCATCTTACATACTGGAATTTATTTTCAATTTAAATTTGCTTA 438
Db 598 ATAATTAACAGGAGCTGAAAAATAGTCTTAATAATAGCCCTTGGGAGGATTTGAGCCCTA 557
Qy 437 ATAGATACAAACGACACCTTAAAAAAGGCTCGCATTTGCATTTGCGGGGTTGAACCCA 378
Db 658 GTTTTGTATTAACAT 673
Qy 377 ATTTTGACATTTGGAAT 362
Db

RESULT 11

PCT-US98-12764-63/c

; Sequence 63, Application PC/TUS9812764

; GENERAL INFORMATION:

; APPLICANT: Gil Choi et. al.

; TITLE OF INVENTION: Borrelia burgdorferi Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 155

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM: Diskette, 3.50 inch, 1.4Mb storage

; MEDIUM TYPE: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US98/12764

; FILING DATE: Herewith

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Brookes, A. Anders

; REGISTRATION NUMBER: 36,373

; REFERENCE/DOCKET NUMBER: PB370PCT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 63:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1761 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

Qy 491 AAAAGCTGATTTTAAGTGGACTTGGGTAAGGGAATGATGTTGGCAGGTGTGTTACTAT 550
Db 3459 AACTGGGCATAT-----AATRAACAACGATAGGAGTAGGCAGAT 3497
Qy 551 GGCTGTGACACAGATTGACAGAAATGTTCTTCCATTTTACATTTGCTAATAATTATACAG 610
Db 3498 TATAGCATCCACACATAGCCTCACTTATTTCCATTTTACATTTGCAATAAACAATGC 3557
Qy 611 GAAGCTGAAAAATAGTCTTAATATAGCCTTGGGAGGATTTGAGCCTAGTTTGTATATTAA 670
Db 3558 AATCTTAAAAAAGACTCGGCATGATATGCGGGTTTGAACCCCAATTTTGATATTGG 3617
Qy 671 CATGGC 677
Db 3618 AATAAGC 3624

Search completed: March 15, 2003, 09:01:10
Job time : 1473.7 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: March 14, 2003, 22:48:52 ; Search time 56.3755 Seconds
(without alignments)
10731.744 Million cell updates/sec

Title: US-09-508-487-20
Perfect score: 862
Sequence: 1 gattttcattggtccag.....atatatttctctgtctgtaa 862

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 501302 seqs, 350932545 residues
Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|---------------------|--------------------|
| 1 | 55 | 6.4 | 416 | 10 | US-09-960-352-4584 | Sequence 4584, Ap |
| 2 | 54.2 | 6.3 | 2000 | 9 | US-09-938-842A-4130 | Sequence 4130, Ap |
| 3 | 51.8 | 6.0 | 640681 | 10 | US-09-790-988-1 | Sequence 1, Appli |
| 4 | 49.2 | 5.7 | 2000 | 9 | US-09-938-842A-5006 | Sequence 5006, Ap |
| 5 | 49.2 | 5.7 | 2002 | 10 | US-09-887-576-4 | Sequence 4, Appli |
| 6 | 48.8 | 5.7 | 2000 | 9 | US-09-938-842A-3659 | Sequence 3659, Ap |
| c 7 | 48.4 | 5.6 | 416 | 10 | US-09-960-352-4584 | Sequence 4584, Ap |
| 8 | 48.2 | 5.6 | 127197 | 9 | US-09-754-853A-1 | Sequence 1, Appli |
| c 9 | 48 | 5.6 | 411 | 10 | US-09-960-352-14521 | Sequence 14521, A |
| 10 | 47.4 | 5.5 | 419 | 10 | US-09-960-352-11234 | Sequence 11234, A |
| c 11 | 47.4 | 5.5 | 640681 | 10 | US-09-790-988-1 | Sequence 1, Appli |
| 12 | 47.2 | 5.5 | 2251 | 10 | US-09-796-256A-11 | Sequence 11, Appli |
| c 13 | 47 | 5.5 | 1058 | 10 | US-09-804-682-30 | Sequence 30, Appli |
| 14 | 46.6 | 5.4 | 431 | 10 | US-09-960-352-5558 | Sequence 5558, Ap |
| c 15 | 46.6 | 5.4 | 2000 | 9 | US-09-938-842A-4502 | Sequence 4502, Ap |
| 16 | 46.2 | 5.4 | 480 | 10 | US-09-960-352-5301 | Sequence 5301, Ap |
| 17 | 46 | 5.3 | 479 | 10 | US-09-960-352-12872 | Sequence 12872, A |
| 18 | 45.8 | 5.3 | 425 | 10 | US-09-960-352-1496 | Sequence 1496, Ap |
| 19 | 45.4 | 5.3 | 446 | 10 | US-09-960-352-3400 | Sequence 3400, Ap |

| | | | | | | |
|------|------|-----|--------|----|---------------------|--------------------|
| c 20 | 45.2 | 5.2 | 2000 | 9 | US-09-938-842A-4747 | Sequence 4747, Ap |
| c 21 | 45.2 | 5.2 | 5006 | 10 | US-09-837-751-7 | Sequence 7, Appli |
| 22 | 45 | 5.2 | 376 | 10 | US-09-960-352-5087 | Sequence 5087, Ap |
| 23 | 44.8 | 5.2 | 516 | 10 | US-09-960-352-5785 | Sequence 5785, Ap |
| 24 | 44.6 | 5.2 | 428 | 10 | US-09-969-373-1172 | Sequence 1172, Ap |
| 25 | 44.6 | 5.2 | 513509 | 9 | US-09-754-853A-4 | Sequence 4, Appli |
| 26 | 44.4 | 5.2 | 413 | 10 | US-09-960-352-2919 | Sequence 2919, Ap |
| 27 | 44.2 | 5.1 | 2846 | 9 | US-10-008-016-1 | Sequence 1, Appli |
| 28 | 44.2 | 5.1 | 4157 | 7 | US-08-556-422-1 | Sequence 1, Appli |
| 29 | 44 | 5.1 | 408 | 10 | US-09-960-352-1221 | Sequence 1221, Ap |
| c 30 | 43.8 | 5.1 | 696 | 10 | US-09-770-693-9 | Sequence 9, Appli |
| c 31 | 43.8 | 5.1 | 73308 | 10 | US-09-954-456-2276 | Sequence 2276, Ap |
| c 32 | 43.6 | 5.1 | 371 | 10 | US-09-969-373-1447 | Sequence 1447, Ap |
| 33 | 43.6 | 5.1 | 414 | 10 | US-09-960-352-6528 | Sequence 6528, Ap |
| c 34 | 43.6 | 5.1 | 419 | 10 | US-09-960-352-11234 | Sequence 11234, A |
| 35 | 43.4 | 5.0 | 2172 | 9 | US-09-966-880A-15 | Sequence 15, Appli |
| 36 | 43.4 | 5.0 | 2818 | 9 | US-09-966-880A-7 | Sequence 7, Appli |
| 37 | 43.4 | 5.0 | 6564 | 9 | US-09-966-880A-10 | Sequence 10, Appli |
| 38 | 43.4 | 5.0 | 11204 | 9 | US-09-966-880A-35 | Sequence 35, Appli |
| c 39 | 43.2 | 5.0 | 299 | 10 | US-09-960-352-13947 | Sequence 13947, A |
| 40 | 43.2 | 5.0 | 385 | 10 | US-09-960-352-1739 | Sequence 1739, Ap |
| c 41 | 43.2 | 5.0 | 26048 | 10 | US-09-764-869-1556 | Sequence 1556, Ap |
| 42 | 43 | 5.0 | 1367 | 9 | US-09-938-842A-4999 | Sequence 4999, Ap |
| 43 | 43 | 5.0 | 9412 | 9 | US-10-032-393-51 | Sequence 51, Appli |
| 44 | 43 | 5.0 | 9412 | 9 | US-10-032-393-52 | Sequence 52, Appli |
| 45 | 43 | 5.0 | 9417 | 9 | US-10-032-393-48 | Sequence 48, Appli |

ALIGNMENTS

RESULT 1
US-09-960-352-4584
; Sequence 4584, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AN
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4584
; LENGTH: 416
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 20-LIB3057-016-Q1-K1-E11
US-09-960-352-4584

| | | | | |
|-----------------------|-----------------|--|-----------|-------------|
| Query Match | 6.4% | Score 55; | DB 10; | Length 416; |
| Best Local Similarity | 49.5% | Pred. No. 0.31; | | |
| Matches 142; | Conservative 0; | Mismatches 145; | Indels 0; | Gaps 0; |
| QY | 21 | AATTTGTGAGATTTTCGACAAATAAGACATTTATTAAGAAGATTCGAAATTCGTAATTTTA 80 | | |
| Db | 129 | AATTTTATAAAAAAATAAATTTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 188 | | |
| QY | 81 | TGGTCAATCAAGAAGCTCTATTGGGAGCGAATTTCAAGTAATCTTTGAAAAAAGTTA 140 | | |
| Db | 189 | TTTAAAAATTAATAATATAAAAAATTAATAAAAAATTAATAAAAAATTAATAAAAAA 248 | | |
| QY | 141 | AATTTAAATAGTTTTTAAAAACCTTTTTTAAATTTTCAATTAATGTACTATAATACCACT 200 | | |
| Db | 249 | AATTTTAAATTTTAAAGAAAAATTTTAAAAATTTTAAAAATTTTAAATTTATAGAACA 308 | | |
| QY | 201 | TTTAAATAAGAGTTTTTATGAATAAATTTTAAATTTTGTGTTGTTGCTAGCCCTTTGTGT 260 | | |
| Db | 309 | TATAAAAAAGATTTTATTAATAAATAATTTAAAAAATAATATGGGTAATTTTATAAAT 368 | | |

Qy 104 GGGAGCGGAATTTCACACTAATACCTTTGAAAAAGTTAAATTAAATAGTTTTAAAAACCT 16

Db 291 TGTGAGTACATGACATTTGTAATAAGTCAAAATACAAAATTTATTTATTTAAACAAAA 350
Qy 164 TTTTAAATTTTCATTAATATGTTACTATAATACAGTTTTTAATAAGAGGTTTTTATGAA 223
Db 351 ACGTTTAGTAAGAATAATATTTATTTTACAAAAAATCTTTGTTTTAAATATTTGTAACAA 410
Qy 224 TAAATTTTAAATGTTGTTGCTAGCCCTTTGTTGTTTTTCTAGCTTTGCTCAAGCTGA 283
Db 411 ATATGTTTATGCTAATTTTAAAGGATTCATTTGAAGTTTGTAGGATTTAAATTTTGATTGA 470
Qy 284 TGATTTCTAAAAGCGCTTTTAAATTTTGGGAGCGGGAGAAAAAATCTTTTAGCTTTATGAACCTAG 343
Db 471 TAAATGTTGATGACAAATTTTAAATAAAAAAGTTTAAAGTAAATTAATTTATTTTGGAAAAAA 530
Qy 344 TAAGAAAGAT 353
Db 531 AATTAATGCT 540

RESULT 5

US-09-887-576-4

Sequence 4, Application US/09887576

Patent No. US2002014047A1

GENERAL INFORMATION:

APPLICANT: Budworth, P.

APPLICANT: Brown, D.

APPLICANT: Chang, H.

APPLICANT: Zhu, T.

APPLICANT: Han, B.

APPLICANT: Wang, X.

APPLICANT: Cooper, Bret

TITLE OF INVENTION: Promoters for regulation of plant expression

FILE REFERENCE: 1360.001US1

CURRENT APPLICATION NUMBER: US/09/887,576

CURRENT FILING DATE: 2001-06-25

PRIOR APPLICATION NUMBER: US 60/213,848

PRIOR FILING DATE: 2000-06-23

PRIOR APPLICATION NUMBER: US 60/214,087

PRIOR FILING DATE: 2000-06-23

PRIOR APPLICATION NUMBER: US 60/258,692

PRIOR FILING DATE: 2000-12-29

NUMBER OF SEQ ID NOS: 875

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 4

LENGTH: 2002

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-09-887-576-4

Query Match 5.7%; Score 49.2; DB 10; Length 2002;
Best Local Similarity 47.4%; Pred. No. 3.5;
Matches 147; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

Qy 44 AAAGACATTTATAAAGAAATTGAAATGCTAATTTTATGGTCAAAATCAAGAGCTCTATT 103
Db 230 AAAGCAAGGATTTAGAAGATTTAAATCTTTTAAATAATTTTTCATTAATAAATAATTA 289
Qy 104 GGGAGCGCAATTTCAAGTAATACITTTGAAAAAGTTAAATTAATAGATTTTAAAAAACC 163
Db 290 TGTGAGTACATGACATTTGTAATAAGTTCAAATACAAAAATTTATTTTAAACAAAAA 349
Qy 164 TTTTAAATTTTCATTAATGTTACTATAATACCAAGTTTAAATAAGAGGTTTTTTATGAA 223
Db 350 ACGTTTAGTAAGAATAATATTTATTTACAAAAAATCTTTGTTTTAAATATTTGTTAAAC 409
Qy 224 TAAATTTTAAATGTTGTTTCTAGCCCTTTGTTGTTTTTCTAGCTTTGCTCAAGCTGA 283
Db 410 ATATGTTTATGCTAATTTTAAAGGATTCATTTGAAGTTTGTAGGATTTAAATTTTGATT 469
Qy 284 TGATTTCTAAAAGCGCTTTTAAATTTTGGGAGCGGGAGAAAAAATCTTTTAGCTTTATGA 343
Db 470 TAAATGTTGATGACAAATTTTAAATAAAAAAGTTTAAAGTAAATTAATTTATTTTGGAAAA 529

Qy 344 TAAGAAAGAT 353
Db 530 AATTAATGCT 539

RESULT 6

US-09-938-842A-3659

Sequence 3659, Application US/09938842A

Patent No. US20020160378A1

GENERAL INFORMATION:

APPLICANT: Harper, Jeff

APPLICANT: Kreps, Joel

APPLICANT: Wang, Xun

APPLICANT: Zhu, Tong

TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINI

TITLE OF INVENTION: SAME, AND METHODS OF USE

FILE REFERENCE: SRIPI300-3

CURRENT APPLICATION NUMBER: US/09/938,842A

CURRENT FILING DATE: 2001-08-24

PRIOR APPLICATION NUMBER: US 60/227,866

PRIOR FILING DATE: 2000-08-24

PRIOR APPLICATION NUMBER: US 60/264,647

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/300,111

PRIOR FILING DATE: 2001-06-22

NUMBER OF SEQ ID NOS: 5379

SEQ ID NO 3659

LENGTH: 2000

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-09-938-842A-3659

Query Match 5.7%; Score 48.8; DB 9; Length 2000;
Best Local Similarity 46.7%; Pred. No. 4.1;
Matches 155; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

Qy 44 AAAGACATTTATAAAGAAATTGAAATGCTAATTTTATGGTCAAAATCAAGAGCTCTATT 103
Db 1486 AAAGCAAGGATTTAGAAGATTTAAATCTTTAATTAATATTTTCATTTGATAAATAATTA 1545
Qy 104 GGGAAAGCAATTTCAAGTAATACTTTGAAAAAGTTTAAATTTTAAATAGTTTTTAAAAAC 163
Db 1546 TGTCAAGTACATGACATTTGTAATAAGTTCAAAATCAAAAATTTATTTATTTTAAACAAA 1605
Qy 164 TTTTAAATTTTCATTAATATGTTACTATAATACCAAGTTTAAATAAGAGGTTTTTTATG 223
Db 1606 ACGTTTAGTAAGAATAATATTTTATTTTAAACAAAAAATCTTTGTTTTAAAAATATTTG 1665
Qy 224 TAAATTTTAAATGTTGTTTGTCTAGCCCTTTTGTGTTTTTCTAGCTTTGCTCAAGCTGA 283
Db 1666 ATATATTTATGCTAATTTTAAAGGATTAATTTGAAGTTTGTGGGATTAATTTGATTGA 1725
Qy 284 TGATTTCTAAAAGCGCTTTTAAATTTTGGGAGCGGGAGAAAAAATCTTTTAGCTTTATGA 343
Db 1726 TAAATGTTGATGACAAATTTTAAATCAAAAAATTTTAAAGTAAATAATTAATTTATTTG 1785
Qy 344 TAAGAAAGATCCTATTTGTCCTCAATTTTATTGA 375
Db 1786 ATTTAATGCTAATGCGCATTAGATTGTAATAA 1817

RESULT 7

US-09-960-352-4584/C

Sequence 4584, Application US/09960352

Patent No. US20020137139A1

GENERAL INFORMATION:

APPLICANT: Warren, Wesley C.

APPLICANT: Tao, Nengbing

APPLICANT: Byatt, John C.

APPLICANT: Machialagan, Nagappan

TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION A

TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION

FILE REFERENCE: 16511.006/37-21(10298)C

GENERAL INFORMATION:
APPLICANT: Kinders, Robert
APPLICANT: Corey, Michael J.
TITLE OF INVENTION: PAL-18 POLYPEPTIDES, NUCLEIC ACIDS
TITLE OF INVENTION: ENCODING THE SAME AND METHODS FOR SCREENING FOR OR
TITLE OF INVENTION: MODULATING THE SAME
FILE REFERENCE: 130001.406
CURRENT APPLICATION NUMBER: US/09/804,682
CURRENT FILING DATE: 2001-03-12
NUMBER OF SEQ ID NOS: 174
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 30
LENGTH: 1058
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7, 9, 88, 462, 505, 531, 582, 586, 607, 624, 658, 679, 680,
LOCATION: 687, 700, 702, 707, 719, 724, 730, 736, 737, 742, 747, 748,
LOCATION: 756, 761, 764, 765, 767, 768, 769, 775, 776, 779, 781, 782,
LOCATION: 785, 787, 788, 792, 793, 794, 796, 799, 800, 807, 817
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc_feature
LOCATION: 818, 819, 821, 822, 823, 828, 829, 830, 832, 837, 838, 849,
LOCATION: 850, 851, 853, 855, 856, 860, 861, 862, 863, 864, 871, 872,
LOCATION: 873, 879, 881, 883, 886, 888, 891, 892, 893, 896, 900, 903,
LOCATION: 904, 905, 906, 909, 913, 914, 916, 919, 921, 925, 929
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc_feature
LOCATION: 930, 931, 950, 951, 954, 957, 963, 965, 966, 967, 968, 969,
LOCATION: 973, 980, 981, 984, 988, 990, 991, 994, 996, 998, 999, 1001,
LOCATION: 1006, 1008, 1009, 1010, 1017, 1020, 1021, 1023, 1024, 1025,
LOCATION: 1027, 1031, 1034, 1035, 1039, 1042, 1043, 1044, 1045
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc_feature
LOCATION: 1046, 1058
OTHER INFORMATION: n = A,T,C or G
US-09-804-682-30

Query Match 5.5%; Score 47; DB 10; Length 1058;
Best Local Similarity 42.6%; Pred. No. 7.8;
Matches 191; Conservative 0; Mismatches 257; Indels 0; Gaps 0;
QY 25 TGTAGATTTTCGACAAATAAGACACATTAATAAGAAATGCTAAATTTATGTT 84
DB 732 TGNAAATNTAGGTTATTGGAAACCAANGNGAAATATTTTCAAGAAAGNAAAGGG 673
QY 85 CAATCAAGAGCTCTATTGGGAGCGAATTCAGTAATACTTTGAAAAAGTTAAAT 144
DB 672 GAATCCAATAACATNTATTGCTATAGAGAAAGAAAGAAATGTAGANTAGGACGTT 613
QY 145 TAAATAGTTTAAAAACCTTTTAAATTCATTAATATGTTACTATATACCAGTTTA 204
DB 612 TTAANCACTGATGGCCCGCTTTTANATANTCTTGATGATATGTTGAATACATGTC 553
QY 205 ATAAAGAGTTTTATGAATAAATTTTAAATGTTGTTTGTGTCAGCTTTTGTGTTT 264
DB 552 AGGATATGAATGAAATTTTCAATTTTAAATATTTTATGCACTATATTTTCAATTTT 493
QY 265 CTAGCTTTGCTCAAGCTGATGATCTTAAAGCGCTTTTAAATTTGGGCGGGAGAAAC 324
DB 492 GTTTTATGTCAGGTTTGAAGAACTTCAAGNAGATATTAATGATACAGATAAATTT 433
QY 325 TTTTAGCTTATGAACCTAGTAAGAAAGATCCTATTGTGCGCATTTTATGAACTTTT 384
DB 432 TCATTTTCAAGAACTTATGTGAACCTGCTTAGTTTATTTTCTTCTTCCCATCT 373
QY 385 TAGGTTTGAATAGGTTCTTTTGTGCTCAAGGAGATATCTGTGGGTTTCTTATCTTG 444
DB 372 TAAGTGCATGCAATTTGTTGATTTGTTTAAATTAATAAAGCTTGACATATATTTTATGT 313
QY 445 GATTGATGCGATGGTATAGGCTTAAT 472

DB 312 AATTGAATTCATTATTGCCCCCTTAAT 285
RESULT 14
US-09-960-352-5558
Sequence 5558, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 5558
LENGTH: 431
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 24-LIB3057-024-Q1-K1-F7
US-09-960-352-5558
Query Match 5.4%; Score 46.6; DB 10; Length 431;
Best Local Similarity 46.0%; Pred. No. 8.4;
Matches 157; Conservative 0; Mismatches 184; Indels 0; Gaps 0;
QY 124 TACTTTGAAAAAGTTAAATTTAAATAGTTTTAAAAACCTTTTTTAAATTTCAATTAAT 183
DB 57 TATTATAATTTGCTTATCTATAAATTTTATTTATTTATTTATTTATTTATTTT 116
QY 184 GTTACTATATATACCAGTTTTTAATAAGAGGTTTTTATGAATAAATTTTAAATTTGTTGTT 243
DB 117 ATTTTATTTTATATATATTTTATTTTATTTTATTTTATTTTATTTTATTTT 176
QY 244 TGCTAGCCTTTTGTGTTTTTCTAGCTTTGCTCAAGCTGATGATCTTAAAGCGCTTTTA 303
DB 177 TTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 236
QY 304 ATTGGGCGGGAGAGAAACTTTTAGCTTATGAAGCTAGTAAGAGATCTTATGTCG 363
DB 237 ATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 296
QY 364 CATTTTATTTGAACCTTTTATTTAGGTTTGAATAGCTTCTTTGCTCAAGGAGATATTC 423
DB 297 TTATTTTATTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 356
QY 424 TTGGGGGTTTCTTATTTCTTTGGATTTGATGACAGTTGCTATA 464
DB 357 TATTATTTTATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 397
RESULT 15
US-09-938-842A-4502/c
Sequence 4502, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCRP1300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 4502
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-4502

Query Match 5.4%; Score 46.6; DB 9; Length 2000;
Best Local Similarity 50.7%; Pred. No. 9.6; Indels 0; Gaps 0;
Matches 112; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
Qy 43 TAAAGACATTATTAAGAAGATTGAAATGCTAATTTTATGGTCAATCAAGAAGCTCTAT 102
Db 1742 TAAAACTTTAGTAGATCTTTTGACTCTTGTATTCAATCAAAACTTTATTAAATTTCTGT 1683
Qy 103 TGGGAAGCGAATTTCAAGTAATACTTTGAAAAAGTTAAATTAATAGTTTTAAAAACC 162
Db 1682 TAAATGTATTATTAGTTTTTGAATTTGTAATTAAGTCTTTTAAAAATTTTCATAAATC 1623
Qy 163 TTTTAAATTTTCATTATATGTTACTATATACCAGTTTAAATAAGAGGTTTTTATGA 222
b 1622 TGAATTTTATGAATCAAAATTTAATAAGTCAATAAATTTTGTGTGTTAATTAA 1563
Qy 223 ATAAATTTTAAATGTTGTTTGTAGCTTTTGTGTTTTT 263
Db 1562 AAAATGTATGATGTTAATGAATTAACATATTGTGTTATT 1522

Search completed: March 15, 2003, 14:21:27
Job time : 525.376 secs

US PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:19:56 ; Search time 31 Seconds
(without alignments)
765.116 Million cell updates/sec

Title: US-09-508-487-21

Perfect score: 904

Sequence: 1 MNRFLIVLLAFVCFSSFAQ.....DINMGQASALGFELSPFKSY 178

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|---------|----------------------|
| 1 | 904 | 100.0 | 178 | AA19965 | Borrelia afzelii A |
| 2 | 786.5 | 87.0 | 177 | AA19965 | Borrelia garinii I |
| 3 | 769.5 | 85.1 | 179 | AA19808 | B. burgdorferi ant |
| 4 | 769.5 | 85.1 | 179 | AA19808 | Borrelia burgdorferi |
| 5 | 687.5 | 76.1 | 158 | AA19809 | B. burgdorferi ant |
| 6 | 160 | 17.7 | 95 | AA19809 | p11 protein sequen |
| 7 | 124 | 13.7 | 25 | AA19809 | Borrelia burgdorferi |
| 8 | 97 | 10.7 | 345 | AA19809 | Corynebacterium gl |
| 9 | 97 | 10.7 | 417 | AA19809 | C glutamicum prote |
| 10 | 95.5 | 10.6 | 479 | AA19809 | B. burgdorferi ant |

| | | | | | | |
|----|------|------|------|----|---------|---------------------|
| 11 | 95.5 | 10.6 | 504 | 20 | AA19964 | B. burgdorferi ant |
| 12 | 85 | 9.4 | 364 | 23 | AB50049 | Listeria monocytog |
| 13 | 83.5 | 9.2 | 572 | 21 | AA19808 | Streptococcus pneu |
| 14 | 83.5 | 9.2 | 650 | 22 | AA19808 | Streptococcus pneu |
| 15 | 82.5 | 9.1 | 215 | 19 | AA19808 | H. pylori GHP132 |
| 16 | 81.5 | 9.0 | 259 | 22 | AA19808 | Human polypeptide |
| 17 | 81.5 | 9.0 | 413 | 22 | AA19808 | Botulinum toxin hea |
| 18 | 81.5 | 9.0 | 848 | 22 | AA19808 | Botulinum toxin hea |
| 19 | 81.5 | 9.0 | 852 | 23 | AB50049 | Botulinum neurotox |
| 20 | 81.5 | 9.0 | 858 | 19 | AA19808 | Recombinant botuli |
| 21 | 81.5 | 9.0 | 1070 | 21 | AA19808 | A manganese supero |
| 22 | 81.5 | 9.0 | 1095 | 21 | AA19808 | A manganese supero |
| 23 | 81.5 | 9.0 | 1169 | 19 | AA19808 | Recombinant botuli |
| 24 | 81.5 | 9.0 | 1291 | 19 | AA19808 | Clostridium botuli |
| 25 | 81 | 9.0 | 305 | 19 | AA19808 | Helicobacter polyp |
| 26 | 79 | 8.7 | 251 | 22 | AA19808 | Micromonospora eve |
| 27 | 78.5 | 8.7 | 663 | 22 | AA19808 | E. coli cellular p |
| 28 | 78.5 | 8.7 | 663 | 22 | AA19808 | Salmonella typhi c |
| 29 | 78.5 | 8.7 | 663 | 22 | AA19808 | E. coli growth and |
| 30 | 77.5 | 8.6 | 444 | 16 | AA19808 | Rhizobium DMP. Rh |
| 31 | 77 | 8.5 | 490 | 23 | AB50049 | Lactococcus lactis |
| 32 | 76.5 | 8.5 | 571 | 22 | AA19808 | Propionibacterium |
| 33 | 76.5 | 8.5 | 635 | 23 | AB50049 | Lactococcus lactis |
| 34 | 76 | 8.4 | 411 | 22 | AA19808 | Staphylococcus aur |
| 35 | 76 | 8.4 | 432 | 22 | AA19808 | Staphylococcus aur |
| 36 | 75.5 | 8.4 | 206 | 22 | AA19808 | Helicobacter pylor |
| 37 | 75.5 | 8.4 | 382 | 23 | AB50049 | Listeria monocytog |
| 38 | 75.5 | 8.4 | 463 | 23 | AB50049 | Listeria monocytog |
| 39 | 75 | 8.3 | 534 | 18 | AA19808 | Arabidopsis thalia |
| 40 | 75 | 8.3 | 534 | 18 | AA19808 | A. thaliana inorga |
| 41 | 75 | 8.3 | 534 | 19 | AA19808 | Arabidopsis thalia |
| 42 | 75 | 8.3 | 534 | 23 | AB50049 | Herbicidally activ |
| 43 | 75 | 8.3 | 535 | 23 | AB50049 | Herbicidally activ |
| 44 | 75 | 8.3 | 538 | 19 | AA19808 | Tomato phosphate t |
| 45 | 74.5 | 8.2 | 180 | 22 | AB50049 | Drosophila melanog |

ALIGNMENTS

RESULT 1
AA19808
ID AA19808 standard; Protein: 178 AA.
AC AA19808;
XX
XX
DT 17-JUN-1999 (first entry)
XX
DE Borrelia afzelii ACAI protein.
XX
KW Borrelia; P13 antigen; spirochaete; vaccine; infection; diagnosis.
XX
OS Borrelia afzelii.
XX
PN WO9912960-A2.
XX
PD 18-MAR-1999.
XX
PF 04-SEP-1998; 98WO-IB01424.
XX
PR 16-SEP-1997; 97US-0059036.
PR 10-SEP-1997; 97DK-0001041.
XX
(SYMB-) SYMBICOM AB.
PA
XX
XX
PI Bergstroem S;
XX
DR WPI; 1999-215027/18.
DR N-PSDB; AAX30099.
XX
XX
PT Nucleic acid from Borrelia burgdorferi encoding virulence associated protein P13
XX

PS Claim 31; Page 111-112; 118pp; English.

XX The present sequence represents a Borrelia afzelii ACAI protein. The
 CC present invention describes an isolated nucleic acid (I) that: (i)
 CC encodes a polypeptide fragment (II) immunologically reactive with
 CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of
 CC Borrelia burgdorferi B313, but not reactive with most proteins from
 CC other spirochaetes; and/or (ii) hybridizes under stringent conditions
 CC to specific nucleic acid sequences but not to genomic DNA from most
 CC other spirochaetes. (I), (II) and transformed cells are useful in
 CC vaccines to protect against infection by B. burgdorferi sensu lato.
 CC (I), (II) and antibodies raised against (II) are used to diagnose such
 CC infections, in standard immunoassays or amplification/hybridization
 CC tests. (I) are also used to produce recombinant (II). The 13 kD protein
 CC is involved in virulence and is highly conserved within B. burgdorferi
 CC sensu lato, but is absent from other Borrelia species (e.g. those
 CC responsible for relapsing fever or avian borreliosis). It should provide
 CC a more specific and sensitive antibody response, and diagnosis,
 CC compared with use of whole bacteria as antigen.

XX Sequence 178 AA;

Query Match 100.0%; Score 904; DB 20; Length 178;
 Best Local Similarity 100.0%; Pred. No. 2.3e-92;
 Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MNKFLIWLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKKDPVFPFLNLFGLFGFI 60
 DB 1 MNKFLIWLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKKDPVFPFLNLFGLFGFI 60
 QY 61 GSAQGDILGGFLLGFDVAGVIGLIITGAYLDIKALDKNAPKAAFKWTWKGMMLAGAVT 120
 DB 61 GSAQGDILGGFLLGFDVAGVIGLIITGAYLDIKALDKNAPKAAFKWTWKGMMLAGAVT 120
 QY 121 MAVTRTEIIIPFTFANSYNRKLKNSLNTAFGGFEPFSDINMGQASALGFELSFKKSY 178
 DB 121 MAVTRTEIIIPFTFANSYNRKLKNSLNTAFGGFEPFSDINMGQASALGFELSFKKSY 178

RESULT 2

AAV04280
 ID AAY04280 standard; Protein; 177 AA.

XX AAY04280;

XX 17-JUN-1999 (first entry)

XX Borrelia garinii IP90 protein.

XX Borrelia; P13 antigen; spirochaete; vaccine; infection; diagnosis.

XX Borrelia garinii.

XX WO912960-A2.

XX 18-MAR-1999.

XX 04-SEP-1998; 98WO-IB01424.

XX 16-SEP-1997; 97US-0059036.

XX 10-SEP-1997; 97DK-0001041.

XX (SYMB-) SYMBICOM AB.

XX Bergstroem S;

XX WPI; 1999-215027/18.

XX N-PSDB; AAX30100.

XX Nucleic acid from Borrelia burgdorferi encoding virulence associated protein P13

XX Claim 31; Page 114-115; 118pp; English.

PS

XX

CC The present sequence represents a Borrelia garinii IP90 protein. The
 CC present invention describes an isolated nucleic acid (I) that: (i)
 CC encodes a polypeptide fragment (II) immunologically reactive with
 CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of
 CC Borrelia burgdorferi B313, but not reactive with most proteins from
 CC other spirochaetes; and/or (ii) hybridizes under stringent conditions
 CC to specific nucleic acid sequences but not to genomic DNA from most
 CC other spirochaetes. (I), (II) and transformed cells are useful in
 CC vaccines to protect against infection by B. burgdorferi sensu lato.
 CC (I), (II) and antibodies raised against (II) are used to diagnose such
 CC infections, in standard immunoassays or amplification/hybridization
 CC tests. (I) are also used to produce recombinant (II). The 13 kD protein
 CC is involved in virulence and is highly conserved within B. burgdorferi
 CC sensu lato, but is absent from other Borrelia species (e.g. those
 CC responsible for relapsing fever or avian borreliosis). It should provide
 CC a more specific and sensitive antibody response, and diagnosis,
 CC compared with use of whole bacteria as antigen.

XX Sequence 177 AA;

Query Match 87.0%; Score 786.5; DB 20; Length 177;
 Best Local Similarity 86.5%; Pred. No. 2.6e-79;
 Matches 154; Conservative 7; Mismatches 16; Indels 1; Gaps 1;
 QY 1 MNKFLIWLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKKDPVFPFLNLFGLFGFI 60
 DB 1 MNKFLIWLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKKDPVFPFLNLFGLFGFI 59
 QY 61 GSAQGDILGGFLLGFDVAGVIGLIITGAYLDIKALDKNAPKAAFKWTWKGMMLAGAVT 120
 DB 60 GSAQGDILGGFLLGFDVAGVIGLIITGAYLDIKALDKNAPKAAFKWTWKGMMLAGAVT 119
 QY 121 MAVTRTEIIIPFTFANSYNRKLKNSLNTAFGGFEPFSDINMGQASALGFELSFKKSY 178
 DB 120 MAVTRTEIIIPFTFANSYNRKLKNSLNTAFGGFEPFSDINMGQASALGFELSFKKSY 177

RESULT 3

AAV19808

ID AAY19808 standard; Protein; 179 AA.

XX AAY19808;

XX 19-JUL-1999 (first entry)

XX B. burgdorferi antigenic protein, f933.aa.

XX Antigenic protein; vaccine; Lyme disease; infection; detection.

XX Borrelia burgdorferi.

XX WO9859071-A1.

XX 30-DEC-1998.

XX 18-JUN-1998; 98WO-US12718.

XX 03-SEP-1997; 97US-0057483.

XX 20-JUN-1997; 97US-0050359.

XX 22-JUL-1997; 97US-0053344.

XX 22-JUL-1997; 97US-0053377.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (MEDI-) MEDIMUNE INC.

XX Choi GH, Erwin AL, Hanson MS, Lathigra R;

XX WPI; 1999-189980/16.

XX N-PSDB; AAX61505.

XX New isolated Borrelia burgdorferi nucleic acids - used to develop products for the diagnosis, prevention and treatment of diseases

PT caused by Borrelia, particularly Lyme disease

PS Claim 12; Page 72; 275pp; English.

XX This sequence represents a Borrelia burgdorferi (Bb) protein of the
CC invention, which is suitable for use in a vaccine. The Bb polypeptides
CC can be used in vaccines for eliciting protective antibodies to members of
CC the Borrelia genus, particularly for the use against Lyme disease in
CC humans and animals. They can be used for preventing or attenuating an
CC infection caused by a member of the Borrelia genus. The products can also
CC be used for detection of members of the Borrelia genus.

XX Sequence 179 AA;

Query Match 85.1%; Score 769.5; DB 20; Length 179;
Best Local Similarity 84.9%; Pred. No. 2e-77;
Matches 152; Conservative 10; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNKFLIYVLLAFVFSFAQADDSKS-AFNLGAGEKLLAYETSKKDPVPPFLNLFGLFGF 59

Db 1 MNKLLIFVLATFCVFSFAQADDSKNGAFGMSAGEKLLVYETSKQDPVPPFLNLFGLFGF 60

QY 60 IGSPAQGDILGGFLLGLFDVAGLILTGAYLDIKALDKNAPKAAFKWTGKGMMLAGAV 119

Db 61 IGSPAQGDILGGFLLGLFDVAGLILTGAYLDIKALDGTITKAAFOQWTGKGMMLAGAV 120

QY 120 TMVTRTEIILPTTFANSYNRKLNLSNIAFGGFEPSFDINMGQASALGFELSKKSY 178

Db 121 TMVTRTEIILPTTFANSYNRKLNLSNIALGFGFEPFDVAMGQSSALGFELSKKSY 179

RESULT 4

AAV04278

ID AAY04278 standard; Protein; 179 AA.

XX AAY04278;

XX 17-JUN-1999 (first entry)

XX Borrelia burgdorferi B31 protein.

XX Borrelia; P13 antigen; spirochaete; vaccine; infection; diagnosis.

XX Borrelia burgdorferi.

XX WO9912960-A2.

XX 18-MAR-1999.

XX 04-SEP-1998; 98WO-IB01424.

XX 16-SEP-1997; 97US-0059036.

XX 10-SEP-1997; 97DK-0001041.

XX (SYMB-) SYMBICOM AB.

XX Bergstroem S;

XX WPI; 1999-215027/18.

XX N-PSDB; AAX30098.

XX Nucleic acid from Borrelia burgdorferi encoding virulence associated
PT protein P13

XX Claim 31; Page 108-109; 118pp; English.

XX The present sequence represents a Borrelia burgdorferi B313 protein.
CC The present invention describes an isolated nucleic acid (I) that: (i)
CC encodes a polypeptide fragment (II) immunologically reactive with
CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of
CC Borrelia burgdorferi B313, but not reactive with most proteins from
CC other spirochaetes; and/or (ii) hybridizes under stringent conditions
CC to specific nucleic acid sequences but not to genomic DNA from most

CC other spirochaetes. (I), (II) and transformed cells are useful in
CC vaccines to protect against infection by B. burgdorferi sensu lato.
CC (I), (II) and antibodies raised against (II) are used to diagnose such
CC infections, in standard immunoassays or amplification/hybridization
CC tests. (I) are also used to produce recombinant (II). The 13 kD protein
CC is involved in virulence and is highly conserved within B. burgdorferi
CC sensu lato, but is absent from other Borrelia species (e.g. those
CC responsible for relapsing fever or avian borreliosis). It should provide
CC a more specific and sensitive antibody response, and diagnosis,
CC compared with use of whole bacteria as antigen.

XX Sequence 179 AA;

Query Match 85.1%; Score 769.5; DB 20; Length 179;
Best Local Similarity 84.9%; Pred. No. 2e-77;
Matches 152; Conservative 10; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNKFLIYVLLAFVFSFAQADDSKS-AFNLGAGEKLLAYETSKKDPVPPFLNLFGLFGF 59

Db 1 MNKLLIFVLATFCVFSFAQADDSKNGAFGMSAGEKLLVYETSKQDPVPPFLNLFGLFGF 60

QY 60 IGSPAQGDILGGFLLGLFDVAGLILTGAYLDIKALDKNAPKAAFKWTGKGMMLAGAV 119

Db 61 IGSPAQGDILGGFLLGLFDVAGLILTGAYLDIKALDGTITKAAFOQWTGKGMMLAGAV 120

QY 120 TMVTRTEIILPTTFANSYNRKLNLSNIAFGGFEPSFDINMGQASALGFELSKKSY 178

Db 121 TMVTRTEIILPTTFANSYNRKLNLSNIALGFGFEPFDVAMGQSSALGFELSKKSY 179

RESULT 5

AAV19809

ID AAV19809 standard; Protein; 158 AA.

XX AAV19809;

XX 19-JUL-1999 (first entry)

XX B. burgdorferi antigenic protein, t933.aa.

XX Antigenic protein; vaccine; Lyme disease; infection; detection.

XX Borrelia burgdorferi.

XX WO9859071-A1.

XX 30-DEC-1998.

XX 18-JUN-1998; 98WO-US12718.

XX 03-SEP-1997; 97US-0057483.

XX 20-JUN-1997; 97US-0050359.

XX 22-JUL-1997; 97US-0053344.

XX 22-JUL-1997; 97US-0053377.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (MEDI-) MEDIMMUNE INC.

XX Choi GH, Erwin AL, Hanson MS, Lathigra R;

XX WPI; 1999-189980/16.

XX N-PSDB; AAX61506.

XX New isolated Borrelia burgdorferi nucleic acids - used to develop
PT products for the diagnosis, prevention and treatment of diseases
PT caused by Borrelia, particularly Lyme disease

XX Claim 12; Page 73; 275pp; English.

XX This sequence represents a Borrelia burgdorferi (Bb) protein of the
CC invention, which is suitable for use in a vaccine. The Bb polypeptides
CC can be used in vaccines for eliciting protective antibodies to members of
CC the Borrelia genus, particularly for the use against Lyme disease in

Best Local Similarity 96.0%; Pred. No. 9.5e-07;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 41 TSKKDPVFPFLNLFGLFGIGSFAQ 65
Db 1 TSKQDPVFPFLNLFGLFGIGSFAQ 25

RESULT 8
ID AAB76781 standard; Protein; 345 AA.
AC AAB76781;
DT 11-APR-2001 (first entry)
DE Corynebacterium glutamicum MCT protein SEQ ID NO:544.
KW Corynebacterium glutamicum; brevibacterium lactofermentum; MCT;
KW membrane construction and membrane transport protein; petroleum spill;
KW hydrocarbon degradation; gram positive aerobic bacterium; marker;
KW identification; microorganism; fine chemical production; transformation;
KW genome mapping; genetic engineering.

OS Corynebacterium glutamicum.
XX WO200100805-A2.
XX 04-JAN-2001.

XX 23-JUN-2000; 2000WO-IB00926.
PR 25-JUN-1999; 99US-0141031.
PR 08-JUL-1999; 99DE-1031454.
PR 08-JUL-1999; 99DE-1031478.
PR 08-JUL-1999; 99DE-1031563.
PR 09-JUL-1999; 99DE-1032122.
PR 09-JUL-1999; 99DE-1032124.
PR 09-JUL-1999; 99DE-1032125.
PR 09-JUL-1999; 99DE-1032128.
PR 09-JUL-1999; 99DE-1032180.
PR 09-JUL-1999; 99DE-1032182.
PR 09-JUL-1999; 99DE-1032190.
PR 09-JUL-1999; 99DE-1032209.
PR 09-JUL-1999; 99DE-1032212.
PR 09-JUL-1999; 99DE-1032227.
PR 09-JUL-1999; 99DE-1032228.
PR 09-JUL-1999; 99DE-1032229.
PR 09-JUL-1999; 99DE-1032230.
PR 14-JUL-1999; 99DE-1032927.
PR 14-JUL-1999; 99DE-1033005.
PR 14-JUL-1999; 99DE-1033006.
PR 27-AUG-1999; 99DE-1040764.
PR 27-AUG-1999; 99DE-1040765.
PR 27-AUG-1999; 99DE-1040766.
PR 27-AUG-1999; 99DE-1040830.
PR 27-AUG-1999; 99DE-1040831.
PR 27-AUG-1999; 99DE-1040832.
PR 27-AUG-1999; 99DE-1040833.
PR 31-AUG-1999; 99DE-1041378.
PR 31-AUG-1999; 99DE-1041379.
PR 31-AUG-1999; 99DE-1041395.
PR 03-SEP-1999; 99DE-1042077.
PR 03-SEP-1999; 99DE-1042078.
PR 03-SEP-1999; 99DE-1042079.
PR 03-SEP-1999; 99DE-1042088.

XX (BADI) BASF AG.
XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Habershauer G;
XX WPI; 2001-071486/08.

DR N-PSDB; AAF68014.
XX Corynebacterium glutamicum nucleic acids encoding membrane construction
PT and membrane transport proteins or their portions, useful for typing or
PT identifying C. glutamicum or related bacteria, and as markers for
PT transformation -
XX
XX PS Claim 20; Page 925-926; 1119pp; English.
XX
CC AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane
CC construction and membrane transport (MCT) proteins given in AAB76510 to
CC AAB76847. The MCT nucleic acids and proteins are useful in the
CC identification of microorganisms which can be used to produce fine
CC chemicals, for modulating fine chemical production in C. glutamicum or
CC related bacteria (e.g. Brevibacterium lactofermentum), the typing or
CC identification of C. glutamicum or related bacteria, as reference points
CC for mapping C. glutamicum genome, and as markers for transformation.
CC AAF68082 and AAF68082 represent sequencing primers which are used in an
CC example from the present invention.
XX
SQ Sequence 345 AA;

Query Match 10.7%; Score 97; DB 22; Length 345;
Best Local Similarity 23.1%; Pred. No. 0.025;
Matches 50; Conservative 24; Mismatches 62; Indels 80; Gaps 9;

QY 5 LIVVLLAFVCVFSFAQADDSKSAFNLGAGEKLLAYETSKK-----DPI 47
Db 63 MTALLLAFC-----LGLGMAVIKSDFLFKVTRLELRVVMKTTTAFVPL 106
QY 48 VP-FLNLLFLGFGIGSFAQGDIL-----GGFLILGFDVIGLITGAYLIDIKALDKNA 100
Db 107 LPLFIQIFLGMCM-----NGLLEIMSAFCKVLILA--VVGTLFLAIQFIAGAVSKKN 160
QY 101 PKAAFKWTNGKMLAGAVTMVTRLTEIIPFTFANSYNRKLNLSNIAFGGFEP---- 156
Db 161 PWKLFFK-----NMLPAYFTALGTSSAAIPVT-----YQOTLKNDVDVNVAGFVPLCA 210
QY 157 -----SFDINMQASALGFEL 172
Db 211 TIHLGSMKIGLFTFAVFMVMDYMEVGLSIGFL 246

RESULT 9
ID AAG93096 standard; Protein; 417 AA.
XX AAG93096;
XX 26-SEP-2001 (first entry)
DE C glutamicum protein fragment SEQ ID NO: 6850.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis.
XX Corynebacterium glutamicum.
XX EF1108790-A2.
XX 20-JUN-2001.
XX 18-DEC-2000; 2000EP-0127688.
XX 16-DEC-1999; 99JP-0377484.
XX 07-APR-2000; 2000JP-0159162.
XX 03-AUG-2000; 2000JP-0280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX Tateishi N, Senoh A, Ikeda M, Ozaki A;

DR WPI: 2001-376931/40.
 XX N-PSDB; AAH68315.
 XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene
 PT
 XX
 PS Claim 17: SEQ ID NO: 6850; 246pp + Sequence Listing; English.
 XX The present invention provides a number of nucleotide and protein
 CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of corynebacterium bacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from Corynebacterium bacterium, and identifying a homologue of a gene derived
 CC from corynebacterium bacterium. Corynebacterium bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described
 CC in the exemplification of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.
 XX
 SQ Sequence 417 AA;

Query Match 10.7%; Score 97; DB 22; Length 417;
 Best Local Similarity 23.1%; Pred. No. 0.032; Length 417;
 Matches 50; Conservative 24; Mismatches 62; Indels 80; Gaps 9;
 QY 5 LIVVLLAFVSSFAQADDSKSAFNLGAGEKLLAYETSKK-----DPI 47
 Db 135 MTALLAFC-----LGLGMAVIKSDTLFKYTRLEERVVMKTIITAFVIPL 178
 QY 48 VP-FLNLFLGFGIGSFAQGDIL-----GGFLILGFDVAGIGLITGAYLDIKALDKNA 100
 Db 179 LFLTFIGFLGGM-----NGGLEIWSAFGKVLILA--VVGPLFLAIQFIAGAVSKKN 232
 QY 101 PKAAFKWTGKGMMLAGAVTMAVTRLTEIIPFTFANSYNRKLNLSLAFGGEP---- 156
 Db 233 PWKLFK-----NMLPAFTALCTSSAATIPVT-----YQQLKNDVDNVNAGVVVPLCA 282
 QY 157 -----SFDINNQSASALGFEL 172
 Db 283 TIHLAGSMKIGLFTFAVVMYDMEVGVGLSIGFL 318

RESULT 10
 AAY19965
 ID AAY19965 standard; Protein: 479 AA.
 AC AAY19965;
 XX

DT 19-JUL-1999 (first entry)
 DE B. burgdorferi antigenic protein, t843.aa.
 KW Antigenic protein; vaccine; Lyme disease; infection; detection.

XX Borrelia burgdorferi.
 OS

XX WO9859071-A1.
 PN

XX 30-DEC-1998.
 PD

XX 18-JUN-1998; 98WO-US12718.
 PF

XX 03-SEP-1997; 97US-0057483.
 PR

XX 20-JUN-1997; 97US-0050359.
 PR

XX 22-JUL-1997; 97US-0053344.
 PR

XX 22-JUL-1997; 97US-0053377.
 PR

XX (HUMA-) HUMAN GENOME SCI INC.
 PA

PA (MEDI-) MEDIMMUNE INC.
 XX Choi GH, Erwin AL, Hanson MS, Lathigra R;
 XX WPI: 1999-189980/16.
 DR N-PSDB; AAX61662.
 XX

PT New isolated Borrelia burgdorferi nucleic acids - used to develop
 PT products for the diagnosis, prevention and treatment of diseases
 PT caused by Borrelia, particularly Lyme disease
 PT
 XX

PS Claim 12: Page 139-140; 275pp; English.
 XX

CC This sequence represents a Borrelia burgdorferi (Bb) protein of the
 CC invention, which is suitable for use in a vaccine. The Bb polypeptides
 CC can be used in vaccines for eliciting protective antibodies to members of
 CC the Borrelia genus, particularly for the use against Lyme disease in
 CC humans and animals. They can be used for preventing or attenuating an
 CC infection caused by a member of the Borrelia genus. The products can also
 CC be used for detection of members of the Borrelia genus.
 XX

SQ Sequence 479 AA;

Query Match 10.6%; Score 95.5; DB 20; Length 479;
 Best Local Similarity 25.5%; Pred. No. 0.056; Length 479;
 Matches 55; Conservative 25; Mismatches 67; Indels 69; Gaps 12;
 QY 6 IVVLLAFVCF-----SSFAQADDSKSAFNLGA-----GEKLLAYETSKKDIPI 47
 Db 93 IIVILPFCVFLAIFFLIWSF---DDLIALSLGLFVFRPEYFGFLYGLNLLPLGLH 149
 QY 48 ----VPF-----LLNLF-----LGFGIGSFAQG--DILGGF-LILGFD 78
 Db 150 SILSPFEFTSLGGVEIVNGDVRGLKNIFYAQLDPSLGKFSFGFAKISSGFYLSIMFG 209
 QY 79 AVGIGLITGAYLDIKALDKNAKPAFAFKWTGKGMMLAGAVTMAVTRLTEIIPFTFANS 138
 Db 210 LPGAAL---GVYKGVHEDKNKVA-----LLFSGALTAFLTGITE---PLEFLFI 254
 QY 139 YNRKLNLSLNTAFGGF-----EPSFDINNQSASALGF 170
 Db 255 FTAPLLYFVHAYSGFALLANFNFTVNTGNSFSTGF 290

RESULT 11
 AAY19964
 ID AAY19964 standard; Protein: 504 AA.
 AC AAY19964;
 XX

DT 19-JUL-1999 (first entry)
 DE B. burgdorferi antigenic protein, f843.aa.
 KW Antigenic protein; vaccine; Lyme disease; infection; detection.

XX Borrelia burgdorferi.
 OS

XX WO9859071-A1.
 PN

XX 30-DEC-1998.
 PD

XX 18-JUN-1998; 98WO-US12718.
 PF

XX 03-SEP-1997; 97US-0057483.
 PR

XX 20-JUN-1997; 97US-0050359.
 PR

XX 22-JUL-1997; 97US-0053344.
 PR

XX 22-JUL-1997; 97US-0053377.
 PR

XX (HUMA-) HUMAN GENOME SCI INC.
 PA

XX (MEDI-) MEDIMMUNE INC.
 PA

PI Choi GH, Erwin AL, Hanson MS, Lathigra R;

XX WPI: 1999-189980/16.
DR N-PSDB; AAX81661.
XX
PT New isolated Borrelia burgdorferi nucleic acids - used to develop
PT products for the diagnosis, prevention and treatment of diseases
PT caused by Borrelia, particularly Lyme disease
XX
PS Claim 12; Page 139; 275pp; English.
XX
CC This sequence represents a Borrelia burgdorferi (Bb) protein of the
CC invention, which is suitable for use in a vaccine. The Bb polypeptides
CC can be used in vaccines for eliciting protective antibodies to members of
CC the Borrelia genus, particularly for the use against Lyme disease in
CC humans and animals. They can be used for preventing or attenuating an
CC infection caused by a member of the Borrelia genus. The products can also
CC be used for detection of members of the Borrelia genus.
XX
SQ Sequence 504 AA;
Query Match 10.6%; Score 95.5; DB 20; Length 504;
Best Local Similarity 25.5%; Pred. No. 0.059;
Matches 55; Conservative 25; Mismatches 67; Indels 69; Gaps 12;
QY 6 IVLLAFVCF-----SFAQADDSKSAFNILGA-----GEKLLAYETSKKDP1--- 47
Db 118 IIVILPCVFLAIFFLIWSSF---DGLIASLGLFVFRFEYFGSLYGLNRLLLPLGLH 174
QY 48 ----VFP-----LLNLF-----LGFIGSFAQG--DIILGGF--LIILGFD 78
Db 175 SILSPFEFTSLGVEIVNGDVTVRGLKNIFYAQLDPSLCKFSGFAKISSGFYLSIMFG 234
QY 79 AVGILGLITGAYLDIKALDKNAPKAFKWTGKGMMLAGAVTMVTRTEIIIPFTFANS 138
Db 235 LPGAAL---GVYKIVHEDKNKVA-----LLFSGALTAFLTGITE---PLEFLFI 279
QY 139 YNRKLNLSIAFGF----EPSPDINMGQASALGF 170
Db 280 FTAPLLVFAHYSGFALLANFENVVIGNSFSTGF 315
RESULT 12
ABBS0049
ID ABB50049 standard; Protein; 364 AA.
AC ABB50049;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes protein #2753.
XX
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.
XX
OS Listeria monocytogenes.
XX
PN WO200177335-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001WO-FR01118.
XX
PR 11-APR-2000; 2000FR-0004629.
XX
PA (JNSP) INST PASTEUR.
XX
PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Maduenio E, De Pablos B, Wehlend J, Kaerst U, Entian K, Hauf J;

PI Rose M, Voss H;
XX WPI: 2002-010914/01.
XX
PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT and prevention of Listeria and related bacterial infections, and
PT related polypeptides -
XX
PS Claim 6; SEQ ID No 2754; 192pp; French.
XX
CC The present invention relates to the genome sequence of Listeria
CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
CC it are useful for selecting probes and primers for detecting genes in L.
CC monocytogenes and related organisms, and for studying genetic
CC polymorphisms and other genomes. The present sequence is a protein
CC encoded by the genome sequence of the present invention. Proteins
CC expressed from the genome sequence are useful for raising specific
CC antibodies, identification of L. monocytogenes and related organisms, and
CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC B12. The genome sequence and proteins encoded by it are also useful for
CC selecting compounds that regulate gene expression and cell replication
CC and modulate L. monocytogenes-related diseases. In addition, the genome
CC sequence and proteins encoded by it are useful in pharmaceutical and
CC vaccines compositions for the treatment or prevention of infections by L.
CC monocytogenes and related organisms.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 364 AA;
Query Match 9.4%; Score 85; DB 23; Length 364;
Best Local Similarity 23.5%; Pred. No. 0.59;
Matches 32; Conservative 26; Mismatches 46; Indels 32; Gaps 6;
QY 37 LAYETSKDPIVPFLLNLFGLGIGSFAQGDILGGFLILGFDVAVGILGLITGAYLDIKAL 96
Db 71 IAYSIADRPGLAPGLLGLIANSIGAGPLGMLGGL-----VGYPVLIVKYLK--- 121
QY 97 DKNAPKAAFKWTGKGM-----LAGAVTMVTRTEIIIPFTFA----NSYNRKLANS 146
Db 122 ----PK-----WAQGLPMPIILSLVGLLMFYVGVPIVWATEMTSFLQGMGS 171
QY 147 LNIAFG---GPEPSFD 159
Db 172 MRFVGVGLGMAAFD 187
RESULT 13
AAY81595
ID AAY81595 standard; Protein; 572 AA.
XX
AC AAY81595;
XX
DT 24-MAY-2000 (first entry)
XX
DE Streptococcus pneumoniae type 4 protein sequence #95.
XX
KW Streptococcus pneumoniae; vaccine; screening; protein antigen;
KW antibacterial; antiinflammatory; meningitis; infection; diagnosis;
KW pneumococcal disease.
XX
OS Streptococcus pneumoniae.
XX
PN WO200006737-A2.
XX
PD 10-FEB-2000.
XX
PF 27-JUL-1999; 99WO-GB02451.
XX
PR 27-JUL-1998; 98GB-0016337.
XX
PR 19-MAR-1999; 99US-0125164.
XX

PA (MICR-) MICROBIAL TECHNIQS LTD.

XX Gilbert CFG, Hansbro PM;

XX WPI; 2000-195300/17.

XX New Streptococcal protein, useful as a vaccine, for diagnosis of
PT pneumococcal diseases and for screening agents capable of antagonizing
PT or inhibiting expression of the protein

XX Claim 1; Page 81; 108pp; English.

XX AAY81501 to AAY81679 represent specifically claimed protein sequences
CC isolated from Streptococcus pneumoniae. AAA05407 to AAA05590 represent
CC specifically claimed nucleotide sequences isolated from S. pneumoniae.
CC The sequences have antibacterial and antiinflammatory properties.
CC The protein sequences, and fragments of them, are useful as immunogens
CC and/or antigens. The nucleotide sequences can be used in vaccines and in
CC diagnostic assays. The proteins and nucleotides can be useful for the
CC detection and diagnosis of S. pneumoniae. The protein sequences are also
CC useful for screening an agent capable of antagonising, inhibiting or
CC interfering with the function or expression of the proteins in which the
CC agent is useful for treatment or prophylaxis of S. pneumoniae infection
CC and meningitis. AAA05591 to AAA05614 represent primers used in the
CC exemplification of the present invention.

XX Sequence 572 AA;

Query Match 9.2%; Score 83.5; DB 21; Length 572;

Best Local Similarity 25.3%; Pred. No. 1.5; Matches 39; Gaps 7;

Matches 39; Conservative 22; Mismatches 54; Indels 39; Gaps 7;

QY 32 AGEKLLAYETSKKDP-----IVPFLNLFGLF-----GIGSFAQGDILGGF 72

Db 341 AGALVLAIRKYYKVPKRSLEGAKSILLPLLTGTLTGFMVLAVNIPMAINTAMNDFLGG- 399

QY 73 LILGFDVAVGIGLITGAYLDIKALDNKAPKAAFKWTGKGMMLA-----GAVTMA----- 122

Db 400 -LGGGSALLGIVLGG-----MMAVDMMGGPVNKAAYVFGTGLTAATVSSGGSVMAAAVMAG 454

QY 123 --VTRLTEIIIPFTFANSYNRKLNKS--LNIAPG 152

Db 455 GMVPPLAIFVATLLFKDKFTKEERNSGLTNIIMG 488

RESULT 14

ID AAU37748

AAU37748 standard; Protein; 650 AA.

XX AC AAU37748;

XX 14-FEB-2002 (first entry)

XX Streptococcus pneumoniae cellular proliferation protein #177.

XX Antisense; prokaryotic cellular proliferation protein;

XX antibiotic; antibacterial; drug design.

XX Streptococcus pneumoniae.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

XX 23-MAY-2000; 2000US-206848P.

XX 26-MAY-2000; 2000US-207727P.

XX 23-OCT-2000; 2000US-242578P.

XX 27-NOV-2000; 2000US-253625P.

XX 22-DEC-2000; 2000US-257931P.

XX 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

DR N-PSDB; AAS55607.

XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids

XX Example 3; Seq ID No 13341; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 650 AA;

Query Match 9.2%; Score 83.5; DB 22; Length 650;

Best Local Similarity 25.3%; Pred. No. 1.8;

Matches 39; Conservative 22; Mismatches 54; Indels 39; Gaps 7;

QY 32 AGEKLLAYETSKKDP-----IVPFLNLFGLF-----GIGSFAQGDILGGF 72

Db 420 AGALVLAIRKYYKVPKRSLEGAKSILLPLLTGTLTGFMVLAVNIPMAINTAMNDFLGG- 478

QY 73 LILGFDVAVGIGLITGAYLDIKALDNKAPKAAFKWTGKGMMLA-----GAVTMA----- 122

Db 479 -LGGGSALLGIVLGG-----MMAVDMMGGPVNKAAYVFGTGLTAATVSSGGSVMAAAVMAG 533

QY 123 --VTRLTEIIIPFTFANSYNRKLNKS--LNIAPG 152

Db 534 GMVPPLAIFVATLLFKDKFTKEERNSGLTNIIMG 567

RESULT 15

AAW98808

ID AAW98808 standard; Protein; 215 AA.

XX AC AAW98808;

XX 31-MAR-1999 (first entry)

XX H. pylori GHPO 1325 protein.

XX GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
KW peptic ulcer disease.

XX Helicobacter pylori.

XX WO9843478-A1.

XX 08-OCT-1998.

XX 01-APR-1998; 98WO-US06371.

XX 29-JUL-1997; 97US-0902615.
PR 01-APR-1997; 97US-0833457.
PR 24-JUN-1997; 97US-0881227.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
XX
PI Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;
XX
DR WPI: 1998-542293/46.
DR N-PSDB; AAX14527.
XX
PT New isolated Helicobacter polynucleotides - used to develop products
PT for the diagnosis, prevention and treatment of Helicobacter
PT infections and gastrointestinal diseases
XX
PS Claim 8; Page 1770-1771; 2054pp; English.
XX
CC This sequence represents a Helicobacter pylori GHPO protein of the
CC invention. The polypeptides can be used for preventing or treating
CC Helicobacter infections, and gastroduodenal diseases associated with
CC these infections, including acute, chronic, and atrophic gastritis, and
CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
CC used for the production of antibodies. The products can also be used for
CC detection and diagnosis.
XX
SQ Sequence 215 AA;

Query Match 9.1%; Score 82.5; DB 19; Length 215;
Best Local Similarity 23.0%; Pred. No. 0.57;
Matches 37; Conservative 18; Mismatches 41; Indels 65; Gaps 6;

QY 5 LIVVLLAFCEVSSFAQADDSKAFN-----LCAGEKLLAY-----ETS 42
DB ::::||||| : :||||| : | : : | :
38 IVMYLMAFCIVSYEDIKNKSAFNIFLLGSLTMAAGLKNVGLNFIGNAAQNFLHA 97
QY 43 KKDPIVP--FLLNLF-----LGFIGSFAQGDILGGFLILGFDVAVG 81
DB ||::| :::| | :||| | | : :|
98 HLDPLIAVLFIYALFYLSHYFFASITAHVSALFALFVIGISHIQGVNLQELSLFLMFLSLG 157
QY 82 IGLILT-----GAYLDIKALDKNAPKAAFKWTWG 110
DB | ||| : | :|
158 IMGILTPYGTGPSTIYYGSGYIQSKDF-----WKWG 188

Search completed: March 14, 2003, 14:24:24
Job time : 33 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:22:16 ; Search time 12.3333 Seconds
(without alignments)
1387.454 Million cell updates/sec

Title: US-09-508-487-21
Perfect score: 904
Sequence: 1 MNKFLIVLLAFCVFSSFAQ.....DINMGQASALGFELSPKSKY 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:.*
1: piri:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------|---------------------|
| 1 | 769.5 | 85.1 | 179 | B70104 | hypothetical prote |
| 2 | 293.5 | 32.5 | 161 | A70207 | conserved hypothet |
| 3 | 285 | 31.5 | 170 | F70243 | conserved hypothet |
| 4 | 240 | 26.5 | 190 | G70239 | conserved hypothet |
| 5 | 95.5 | 10.6 | 574 | D70114 | PTS system, maltos |
| 6 | 89.5 | 9.9 | 469 | H97064 | probable sugar-pro |
| 7 | 88 | 9.7 | 312 | F75346 | phospho-N-acetylmu |
| 8 | 87.5 | 9.7 | 616 | G98325 | probable c4-dicarb |
| 9 | 87.5 | 9.7 | 616 | A22957 | hypothetical prote |
| 10 | 86.5 | 9.6 | 545 | S59143 | NADH2 dehydrogenas |
| 11 | 85.5 | 9.5 | 482 | F71969 | hypothetical prote |
| 12 | 85 | 9.4 | 363 | AE1712 | PTS system, fructo |
| 13 | 85 | 9.4 | 364 | AG1341 | PTS system, fructo |
| 14 | 85 | 9.4 | 538 | A41901 | metabolite export |
| 15 | 84.5 | 9.3 | 674 | T19495 | hypothetical prote |
| 16 | 83.5 | 9.2 | 449 | S37867 | hypothetical prote |
| 17 | 83.5 | 9.2 | 650 | C95101 | PTS system, fructo |
| 18 | 83.5 | 9.2 | 650 | D79269 | phosphotransferase |
| 19 | 83.5 | 9.2 | 709 | S75212 | comE protein - Syn |
| 20 | 83 | 9.2 | 445 | B83885 | cytosolic permease |
| 21 | 82.5 | 9.1 | 559 | A45620 | cyto villin homolog |
| 22 | 82.5 | 9.1 | 559 | S49143 | EG10 protein - tap |
| 23 | 81.5 | 9.0 | 692 | H70362 | glycogen phosphory |
| 24 | 81.5 | 9.0 | 1291 | A48940 | bontoxilysin (EC 3 |
| 25 | 81 | 9.0 | 214 | S76538 | hypothetical prote |
| 26 | 81 | 9.0 | 305 | F64599 | outer membrane pro |
| 27 | 81 | 9.0 | 507 | B69957 | conserved hypothet |
| 28 | 80.5 | 8.9 | 357 | H70346 | undecaprenyl-phosp |
| 29 | 80.5 | 8.9 | 371 | S55588 | D-nopaline dehydro |

| | | | | | | |
|----|------|-----|------|---|--------|--------------------|
| 30 | 80.5 | 8.9 | 882 | 2 | A42855 | N-heparan sulfate |
| 31 | 80 | 8.8 | 183 | 2 | H72250 | conserved hypothet |
| 32 | 80 | 8.8 | 336 | 1 | S75272 | cytochrome d ubiqu |
| 33 | 80 | 8.8 | 779 | 2 | AD1457 | ATP dependent heli |
| 34 | 79.5 | 8.8 | 419 | 2 | A47033 | chloramphenicol re |
| 35 | 79.5 | 8.8 | 419 | 2 | JQ1201 | CmlA protein - pse |
| 36 | 79 | 8.7 | 429 | 2 | B83723 | hypothetical prote |
| 37 | 79 | 8.7 | 539 | 2 | T02512 | hypothetical prote |
| 38 | 78.5 | 8.7 | 372 | 2 | AB0276 | probable membrane |
| 39 | 78.5 | 8.7 | 436 | 2 | D83498 | C4-dicarboxylate t |
| 40 | 78.5 | 8.7 | 663 | 2 | A85540 | bo-type ubiquinol |
| 41 | 78.5 | 8.7 | 663 | 2 | E90689 | cytochrome o ubiqu |
| 42 | 78.5 | 8.7 | 663 | 2 | AE0557 | cytochrome o ubiqu |
| 43 | 78.5 | 8.7 | 663 | 2 | B42226 | bo-type ubiquinol |
| 44 | 78.5 | 8.7 | 1291 | 2 | I40631 | non-proteolytic bo |
| 45 | 78 | 8.6 | 235 | 2 | D69597 | probable holocyto |

ALIGNMENTS

RESULT 1

B70104
hypothetical protein BB0034 - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Mar-2000
C:Accession: B70104
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; W
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; V
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: B70104
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-179 <K1E>
A:Cross-references: GB:AE001117; GB:AE000783; NID:92687907; PIDN:AAC66426.1; PID:922
A:Experimental source: strain B31
C:Superfamily: Lyme disease spirochete plasmid hypothetical protein BBA01

| | | | | | | | | | |
|-----------------------|-------|--------------|-----------|------------|----------------------------------|---------------|------|------|----|
| Query Match | 85.1% | Score | 769.5; | DB 2; | Length | 179; | | | |
| Best Local Similarity | 84.9% | Pred. No. | 2.8e-63; | | | | | | |
| Matches | 152; | Conservative | 10; | Mismatches | 16; | Indels | 1; | Gaps | 1; |
| QY | 1 | MNKFLLIVLLAF | CVFSSFAQ | ADSKS- | AFNLGAGEKLLAYETSKKDP | IVPFLNLF | LGFG | 59 | |
| | | | | | | | | | |
| Db | 1 | MNKLIFVLAT | FCVFSFAQ | ADSKS | AGKLGAFMSAGEKLLVYETSKQDP | IVPFLNLF | LGFG | 60 | |
| QY | 60 | IGSFAQGD | IIGGFLILG | FDVAVG | IGLILTGAYLIDIKALDKNAPKAFK | TKTWGKMMLAGAV | 119 | | |
| | | | | | | | | | |
| Db | 61 | IGSFAQGD | IIGGSLILG | FDVAVG | IGLILAGAYLIDIKALDKGITKKAFAQ | FTWKGVMLAGV | 120 | | |
| QY | 120 | TMAVTRLE | TEIIPFTF | FANSYNRKLN | SNIAFGGFEPSFDINMGQASALGFELSPKSKY | 178 | | | |
| | | | | | | | | | |
| Db | 121 | TMAVTRLE | TEIIPFTF | FANSYNRKLN | SNIAFGGFEPSFDVAMGQSSALGFELSPKSKY | 179 | | | |

RESULT 2

A70207
conserved hypothetical protein BBA01 - Lyme disease spirochete plasmid A/lp54
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Mar-2000
C:Accession: A70207
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; W
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; V
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: A70207

A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-161 <KLE>
A:Cross-references: GB:AE000790; NID:g2690224; PIDN:AAC66226.1; PID:g2690225; TIGR:BBA01
A:Experimental source: strain B31
C:Genetics:
A:Genome: plasmid
C:Superfamily: Lyme disease spirochete plasmid hypothetical protein BBA01

Query Match 32.5%; Score 293.5; DB 2; Length 161;
Best Local Similarity 41.1%; Pred. No. 9.6e-20;
Matches 74; Conservative 24; Mismatches 61; Indels 21; Gaps 6;

QY 1 MNKEFLIVLLAFCAVSSFAQAADSKSAFNAGAGEKKLAYETSKKDPIVPFLLNLFLGFGI 60
| | : : : | | : : : | | : : : | | : : : | | : : : | | : : : |
Db 1 MKKIITLILFSLTIQIFA-TKTQNRIEKG-I-ESEFNKYDKREKNPISGPPELNLFDPFI 58
| | : : : | | : : : | | : : : | | : : : | | : : : | | : : : |

QY 61 GSFAQGDI LGGFLILGDPAVGIGILTGA YLDIKALDKNAPKA AFKWTKWGMMLAG-AV 119
||| ||| ||| : ||| : | | : | | : ||| : ||| : | | : ||| : ||| :
Db 59 GSFVQGVIGGGVGLGNLLGAILWGTFIIN-----HRETQLTGYILIGVGA 106
| | : : : | | : : : | | : : : | | : : : | | : : : | | : : : |

QY 120 TMAVTRLT-ELIIPTFANSYNRKLNLSNIARFGPEPSFDINMGQASALGFELSPKSY 178
| | : | | : ||||||| : | | : | | : ||| : ||| : | | : ||| : ||| :
Db 107 SMLVTSTVSLLIPTFFANRHNNENLKRLSAAELAGFEPNFDLGIN----GFOLSFKKSY 161
| | : | | : ||||||| : | | : | | : ||| : ||| : | | : ||| : ||| :

RESULT 3
F70243
conserved hypothetical protein BB131 - Lyme disease spirochete plasmid I/lp28-4
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Mar-2000
C:Accession: F70243
R:Fraser, C.M.; Castjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; White-son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.J.; Salzberg, S.; Hanson, M.; Vugt, B.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; PMID:98065943; PMID:9403685
A:Accession: F70243
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-170 <KLE>
A:Cross-references: GB:AE000789; NID:g2690079; PIDN:AAC66194.1; PID:g2690091; TIGR:BB131
A:Experimental source: strain B31
C:Genetics:
A:Genome: plasmid
C:Superfamily: Lyme disease spirochete plasmid hypothetical protein BBA01

Query Match 31.5%; Score 285; DB 2; Length 170;
Best Local Similarity 38.8%; Pred. No. 6.1e-19;
Matches 69; Conservative 25; Mismatches 70; Indels 14; Gaps 3;

QY 1 MNKEFLIVLLAFCAVSSFAQAADSKSAFNAGAGEKKLAYETSKKDPIVPFLLNLFLGFGI 60
| | : : : | | : : : | | : : : | | : : : | | : : : | | : : : |
Db 7 MKKIITLILFSLTIQIFATQD--KLEKS VGSIEFTINKYSEKATILAPELLNLFLTGI 64
| | : : : | | : : : | | : : : | | : : : | | : : : | | : : : |

QY 61 GSFAQGDI LGGFLILGDPAVGIGILTGA YLDIKALDKNAPKA AFKWTKWGMMLAGAVT 120
||| ||| ||| : ||| : | | : | | : ||| : ||| : | | : ||| : ||| :
Db 65 GSFVQGVIGGGVAVGSQLGGILCTAGNLTGHGTDDTRA-----TTGHIITTVIGVT 117
| | : : : | | : : : | | : : : | | : : : | | : : : | | : : : |

QY 121 MAVTRLTEIIPTFANSYNRKLNLSNIARFGPEPSFDINMGQASALGFELSPKSY 178
| | : | | : ||||||| : | | : | | : ||| : ||| : | | : ||| : ||| :
Db 118 IIASHIASLIPTFFANKHNANUKRLGIDIA GFEPNFDIGS-----GFOLSFKKRY 170
| | : | | : ||||||| : | | : | | : ||| : ||| : | | : ||| : ||| :

RESULT 4
G70239
conserved hypothetical protein BBH41 - Lyme disease spirochete plasmid H/lp28-3
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Mar-2000
C:Accession: G70239
R:Fraser, C.M.; Castjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White-

[illegible]

THE UNIVERSITY OF CHICAGO PRESS

A; Map position: linear chromosome

QY 38 AYETSKKDPIVPFLLNULFGLFGIGSFAQGDILGGFLILG----FDAVGIGLIITGAYLDI 93

QY 38 AYETSKKDPIVPFLLNULFGLFGIGSFAQGDILGGFLILG----FDAVGIGLIITGAYLDI 93

```

Query Match          9.7%; Score 87.5; DB 2; Length 616;
Best Local Similarity 22.0%; Pred. No. 2.8;
Matches 42; Conservative 35; Mismatches 71; Indels 43; Gaps 9;

          Y          Y          Y          Y          Y          Y          Y          Y          Y          Y
9  LLAFCVSSFAQADDSKSAFNLG-----AGEKLAVETSKKQPIV----- 49
      ||  ||  :  :  :  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
301 LTVSLVFSGAGSSVANAFAAGASTFQPELVKHYPPAQAGAITAA--TSVLDNVIPSTA 358
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
50  -PLLNLFTGFGIGSQAQDILGFLIIFDVGIGLITGAYLIDIKALDNKAPKA--APK 106
      :  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  :
359 FLIATATNLSVGS|--LVGGFFAGGLMAVCLGVAI-----HLSVRSVD--TLPRATGAER 410
      :  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  :
107 W-----TWKGMMLAGAYTMVATRIITLIPFFFANSYNRKLKLSNLNIARGGPEPSF 158
      :  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  :
411 WRSATAATPAFLGVVVVVVVVIRIGIVITTEAA---ALAAALYTLILGFGYRLGVGRIFATF 467
      :  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  :
159 DINMGQASALG 169
      :  ||  ||  ||
468 ROSAGAAAAIG 478
      :  ||  ||  ||

```

RESULT 10
S59143
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - land snail mitochondrion
C:Species: Mitochondrion Albinaria coerulea (land snail)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 03-Jun-2002
C:Accession: S59143
R:Hatzoglou, E.; Rodakis, G.C.; Lecanidou, R.
Genetics 140, 1353-1366, 1995
A:Title: Complete sequence and gene organization of the mitochondrial genome of the land
A:Reference number: S59143; MUID:96120351; PMID:7498775
A:Accession: S59143
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-545 <HAT>
A:Cross-references: EMBL:X83390; NID:g975668; PIDN:CAA58296.1; PID:g975669; GSPDB:GN0013
C:Genetics:
A:Gene: ND5
A:Genome: mitochondrion
A:Genetic code: SGC4
A:Start codon: ATT
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 5
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

| Query Match | 9.6% | Score 86.5 | DB 2 | Length 545 |
|-----------------------|-------|--|---------------|------------|
| Best Local Similarity | 29.1% | Pred. No. 3.1 | | |
| Matches | 39 | Conservative 21 | Mismatches 37 | Indels 37 |
| Gaps | | | | |
| Qy | 4 | FLIVLL-ACVSVSSFAQADSKSAFNLGAGEKLLAYETSKKQPIVFLLNLPLGFGIGS | 62 | |
| Db | 51 | FLVVLIISSCVF-LFANEYMSDEHYNRFGWILISVAS-----MGILILS | 96 | |
| Qy | 63 | FAQGDILGGFLILGFDVAVGIGLILTGAYLDIKALDKNAPRAAFKWTGCKMMLAGAVTMA | 129 | |
| Db | 97 | --GSIF--LILLGWDGLGILTSFALIAYYD-----NYNASSAF-----LTLM | 134 | |
| Qy | 123 | VTRLTEIILPPTFA | 136 | |
| Db | 135 | TNRLGDVLIATFS | 148 | |

RESULT 11
F71969
hypothetical protein jhp0131 - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 20-Aug-1999
C:Accession: F71969
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human
A:Reference number: A71800; MUID:99120557; PMID:9923682

```

A:Accession: F71969
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-482 <ARN>
A:Cross-references: GB:AE001452; GB:AE001439; NID:g4154639; PIDN:AAD05710.1; PI:
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0131
C:Superfamily: 2-oxoglutarate/malate translocator

Query Match          9.5%; Score 85.5; DB 2; Length 482;
Best Local Similarity 24.2%; Pred. No. 3.3;
Matches 39; Conservative 18; Mismatches 39; Indels 65; Gaps 8;

Qy  5  LIVVLLAFCVSSFAQADDSKAEN-----LGAGEKLLAY-----ETS 42
Db  305 IVVWLMAFCIVSYEDIINKSAFNILLGLSLUTMAGGLKNVGFNLFGNAQNFLFHA 364
                                     : : : : :
Qy  43  KKDPVLP--FLNLNLF-----LGFGIGSFAQG---DIUGGFLILGFD 78
                                     : : : : :
Db  365 HLDPLIAVLFIVALFYLSHYFFASITAHVSALFALFVGIGSHIOGVNLQELSLFMLS LG 424
                                     : : : : :

Qy  79  AVGI-----GILTGA-YLDIKALDKNAPKAARFTWG 110
Db  425 I MGILTPGTGPSTIYVGGYIOSKDF-----WKWG 455

```

RESULT 12
 AE1712
 PTS system, fructose-specific enzyme IIC component homolog lin2240 [imported]
 C:Species: *Listeria innocua*
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AE1712
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.;
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.;
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournon,
 ok, C.; Schluerer, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.;
 A:Title: Comparative genomics of *Listeria* species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AE1712
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-363 <GLA>
 A:Cross-references: GB:AL592022; PIDN:CA97469.1; PID:gl6414753; GSPDB:GN00178;
 A:Experimental source: strain Cl1p1262
 C:Genetics:
 A:Gene: lin2240

| | | | | | |
|----|-----------------------|---|----------------|------------|-------------|
| | Query Match | 9.4%; | Score 85; | DB 2; | Length 363; |
| | Best Local Similarity | 23.5%; | Pred. No. 2.8; | | |
| | Matches 32; | Conservative 26; | Mismatches 46; | Indels 32; | Gaps 6; |
| Qy | 37 | LAYETSKDPIVPLLNFLFLGFGTSGFAQGDIIGGFLILGFDAVGIGLILTGAVLDIKAL | 96 | | |
| Dd | 71 | IAYSTADPGTAGLLMLGLIANSICAGPLGGMLGGYL-----VGYFVLIVLVKLV--- | 121 | | |
| Qy | 97 | DKNAPKAFTWTKGMM-----LAGAVTMAVTRLTETIIIPFFA-----NSYNRKLKNS | 146 | | |
| Dd | 122 | ---PK-----WAQGLPMMIIPLISSLVIGLLMYFVGVPYVWATEAMTSFLQQMGQS | 171 | | |
| Qy | 147 | LNIAFG----GPESFD | 159 | | |
| Dd | 172 | MRVFVGAVLGAAAFD | 187 | | |

```

RESULT 13
AG1341
PTS system, fructose-specific enzyme IIC component homolog lmo2135 [imported] -
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AG1341

```

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria species*.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1341
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-364 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAD00213.1; PID:g16411605; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo2135

Query Match 9.4%; Score 85; DB 2; Length 364;
Best Local Similarity 23.5%; Pred. No. 2.8;
Matches 32; Conservative 26; Mismatches 46; Indels 32; Gaps 6;

Qy 37 LAYETSKDPIVPELNLFLGFGIGSFAQGDILGFLILGFDVAVGIGLILTCAYLIDIKAL 96
Db 71 IAYSTADRPAGIAPGLLMGLIANSGAGFLGMLGGYL-----VGYFVLILVKLVK--- 121
Qy 97 DKNAPKAAFKWTGKGM-----LAGAVTMAVTRLTEIIPFFA-----NSYNRKLKNS 146
Db 122 ----PK-----WAQGLPMWMIIPLISSLVGLLMFYVGVPIVWATEAMTSFLQMGGS 171
Qy 147 LNIAGF---GPEPSFD 159
Db 172 MRFVFGAVLGAMAAFD 187

RESULT 14

A41901
metabolite export pump of tetracycline C resistance TcmA - Streptomyces glaucescens
C:Species: Streptomyces glaucescens
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 22-Oct-1999
C:Accession: A41901; S27687
R:Guilfoile, P.G.; Hutchinson, C.R.
J. Bacteriol. 174, 3651-3658, 1992
A:Title: Sequence and transcriptional analysis of the Streptomyces glaucescens tcmAR tet
A:Reference number: A41901; MUID:92276347; PMID:1592819
A:Contents: GLA.O, ETH 22794
A:Accession: A41901
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-538 <GUI>
A:Cross-references: EMBL:M80674; NID:g153488; PIDN:AAA67509.1; PID:g153490
A:Note: sequence extracted from NCBI backbone (NCBIP:104630)
C:Superfamily: lincomycin-resistance protein lmrB
C:Keywords: transmembrane protein

Query Match 9.4%; Score 85; DB 2; Length 538;
Best Local Similarity 24.6%; Pred. No. 4.2;
Matches 33; Conservative 19; Mismatches 38; Indels 44; Gaps 6;

Qy 49 PELLNLFLGFGIG-----SFAQGDILGFLI--LGFD-----VGIILITGAYLIDIKAL 95
Db 145 PGLNMAIGISGVVVGAAAGPIIGLLVQHVGVAVFFINVPVGLAALVAGLILTDAL 204
Qy 96 LDKNAPKA-----AFKWTGK-----GMMLAGAV-----TMAVT 124
Db 205 RAERAPKSFVSGIVLLSGAMFCLVWGLIKAPMGWGDRLTGLFLAAAVLAFAGFTLRES 264
Qy 125 RLTEIIPFFANS 138
Db 265 RATEPLMPLAMERS 278

RESULT 15

T19495

hypothetical protein C27A7.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T19495
R:Harris, B.
submitted to the EMBL Data Library, October 1996
A:Reference number: T19132
A:Accession: T19495
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-674 <WIL>
A:Cross-references: EMBL:Z81041; PIDN:CA802785.1; GSPDB:GN00023; CESP:C27A7.3
A:Experimental source: clone C27A7
C:Genetics:
A:Gene: CESP:C27A7.3
A:Map position: 5
A:Introns: 61/1; 98/3; 191/1; 220/3; 268/3; 340/1; 392/2; 475/2; 496/3; 537/1; 588/1

Query Match 9.3%; Score 84.5; DB 2; Length 674;
Best Local Similarity 24.0%; Pred. No. 5.8;
Matches 37; Conservative 21; Mismatches 47; Indels 49; Gaps 8;

Qy 4 FLIVVLLAFCVFSSFAQADDSKAFNLGAGEKLLAYETSKKDPVIVPFLNLFLGFGIGSF 63
Db 41 FLFILLLCFSVSFSFTLI-----IKNYQT-----FLISFFFRIVWSL 78
Qy 64 AQGDILGGFLILGFDVAVGIGLILTCAYLIDIKALDKNAPKAAFKWTGKGMMLAGAVTMAV 123
Db 79 AMLSFIAGLVLGLVAAA---TISGS-----KNLPTAEYKW-----AGCENLGK 119
Qy 124 TRLTE-----IIIPFT-FANSY-NRKLKNSLNI 149
Db 120 QCIDGFSFPLVILSFDGFAKEYLERRIVKSL 153

Search completed: March 14, 2003, 14:25:41
Job time : 14.3333 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:25:51 ; Search time 10.0562 seconds
(without alignments)
820.439 Million cell updates/sec

Title: US-09-508-487-19
Perfect score: 902
Sequence: 1 MNKLIFVLATCFVFSFAQ.....DVAMGSSALGFELSFKKSX 179

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pbp.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pbp.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pbp.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pbp.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pbp.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pbp.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pbp.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pbp.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pbp.*
- 10: /cgn2_6/ptodata/2/pubpaa/US03_PUBCOMB.pbp.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pbp.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pbp.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pbp.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pbp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 84.5 | 9.4 | 378 | 9 | US-10-260-877-48 |
| 2 | 83.5 | 9.3 | 384 | 9 | US-09-738-626-3606 |
| 3 | 80.5 | 8.9 | 650 | 10 | US-09-815-242-13341 |
| 4 | 80 | 8.9 | 487 | 9 | US-09-738-626-4875 |
| 5 | 79.5 | 8.8 | 319 | 10 | US-09-815-242-13924 |
| 6 | 79.5 | 8.8 | 417 | 9 | US-09-738-626-6850 |
| 7 | 79 | 8.8 | 346 | 9 | US-09-738-626-5811 |
| 8 | 78.5 | 8.7 | 513 | 9 | US-09-738-626-6860 |
| 9 | 77.5 | 8.6 | 513 | 9 | US-10-051-902-20 |
| 10 | 77.5 | 8.6 | 513 | 9 | US-10-051-909-20 |
| 11 | 74.5 | 8.3 | 570 | 10 | US-09-815-242-11951 |
| 12 | 74.5 | 8.3 | 708 | 9 | US-09-738-626-5024 |
| 13 | 74 | 8.2 | 417 | 10 | US-09-815-242-10165 |
| 14 | 73 | 8.1 | 451 | 10 | US-09-741-669-422 |
| 15 | 72.5 | 8.0 | 345 | 9 | US-09-738-626-4062 |
| 16 | 72.5 | 8.0 | 468 | 10 | US-09-815-242-5144 |
| 17 | 72.5 | 8.0 | 693 | 9 | US-09-365-536A-14 |
| 18 | 72.5 | 8.0 | 1161 | 8 | US-08-910-386A-20 |
| 19 | 72.5 | 8.0 | 1294 | 9 | US-09-836-499-2 |

| | | | | | |
|----|------|-----|------|----|---------------------|
| 20 | 72.5 | 8.0 | 1309 | 9 | US-09-836-499-5 |
| 21 | 72 | 8.0 | 962 | 9 | US-09-738-626-3800 |
| 22 | 71.5 | 7.9 | 255 | 10 | US-09-881-752A-106 |
| 23 | 71.5 | 7.9 | 413 | 9 | US-09-910-186A-22 |
| 24 | 71.5 | 7.9 | 544 | 9 | US-09-738-626-3927 |
| 25 | 71.5 | 7.9 | 544 | 10 | US-09-948-777-2 |
| 26 | 71.5 | 7.9 | 852 | 9 | US-10-011-588-25 |
| 27 | 71 | 7.9 | 225 | 9 | US-09-860-670-109 |
| 28 | 71 | 7.9 | 257 | 9 | US-09-738-626-3912 |
| 29 | 71 | 7.9 | 266 | 9 | US-10-131-406-7 |
| 30 | 71 | 7.9 | 380 | 10 | US-09-815-242-10609 |
| 31 | 71 | 7.9 | 431 | 10 | US-09-815-242-11033 |
| 32 | 71 | 7.9 | 670 | 9 | US-10-060-763-10 |
| 33 | 71 | 7.9 | 670 | 12 | US-10-063-763-10 |
| 34 | 71 | 7.9 | 702 | 10 | US-09-907-479-2 |
| 35 | 70.5 | 7.8 | 319 | 10 | US-09-741-669-379 |
| 36 | 70.5 | 7.8 | 319 | 10 | US-09-815-242-10398 |
| 37 | 70.5 | 7.8 | 436 | 10 | US-09-815-242-13218 |
| 38 | 70.5 | 7.8 | 712 | 9 | US-10-095-139-2 |
| 39 | 70.5 | 7.8 | 712 | 9 | US-10-095-139-13 |
| 40 | 70.5 | 7.8 | 712 | 9 | US-10-095-139-23 |
| 41 | 70 | 7.8 | 210 | 10 | US-09-764-855-90 |
| 42 | 70 | 7.8 | 321 | 9 | US-09-738-626-3676 |
| 43 | 70 | 7.8 | 411 | 10 | US-09-815-242-5875 |
| 44 | 70 | 7.8 | 432 | 10 | US-09-815-242-13058 |
| 45 | 69 | 7.6 | 268 | 10 | US-09-815-242-12057 |

ALIGNMENTS

RESULT 1

US-10-260-877-48

; Sequence 48, Application US/10260877

; Publication No. US20030021813A1

; GENERAL INFORMATION:

; APPLICANT: Abbott Laboratories

; APPLICANT: Chovan, Linda E.

; APPLICANT: Hessler, Paul E.

; APPLICANT: Reich, Karl A.

; TITLE OF INVENTION: ESSENTIAL BACTERIA GENES AND GENOME

; TITLE OF INVENTION: SCANNING IN HAEMOPHILUS INFLUENZAE FOR THE IDENTIFICATION OF

; FILE REFERENCE: 'ESSENTIAL GENES'

; FILE REFERENCE: 6565 US.P1

; CURRENT APPLICATION NUMBER: US/10/260,877

; CURRENT FILING DATE: 2002-09-30

; PRIOR APPLICATION NUMBER: US/09/649,145

; PRIOR FILING DATE: 2000-08-25

; NUMBER OF SEQ ID NOS: 137

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 48

; LENGTH: 378

; TYPE: PRT

; ORGANISM: H. influenzae

US-10-260-877-48

Query Match 9.4% Score 84.5; DB 9; Length 378;

Best Local Similarity 28.3%; Pred. No. 0.44; Mismatches 19; Gaps 5;

Matches 34; Conservative 19; Indels 19; Gaps 5;

QY 14 VFSSFAQANDSKNGAFMGMSAGEKLLVYETSQDPIVFPF-----LLNLFGLFGIGSFAQG 67

Db 73 VKSFGLSGDGKTAIEMAAASGLHLVPEKRNPLTTSYGTGELIKLALDLGVESFTLG 132

QY 68 DILGSSILGPDVAVGILIGLAGYLDIKAGFQWTKGVMLAGVVTMVTFL 127

Db 133 --IGGS---ATNDGGVGMLQA---LGMQCLDSQDKPIGF-----GGAELANIVKIDVQOL 179

RESULT 2

US-09-738-626-3606

; Sequence 3606, Application US/09738626

; Publication No. US20020197605A1

```
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3606
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3606
```

Query Match 9.3%; Score 83.5; DB 9; Length 384;

Best Local Similarity 24.3%; Pred. No. 0.56; Mismatches 24; Indels 59; Gaps 12;

Matches 56; Conservative 24; Mismatches 24; Indels 59; Gaps 12;

```
QY 6 IFVATFCVFSFAQAANDSK-----NGAFGMSA----- 33
: : : : : : : : : : : : : : : : : : : : : : : : :
Db 9 VFVGVILLFTAGWAANHFASVLVIREQDLYSSVLVNGAFGIYALGLPSLLAGGVLD 68
: : : : : : : : : : : : : : : : : : : : : : : : :
QY 34 --GEKLLVYETSKQDPVIFLNLFGFGIGSFAQGDILGSLILGFDVAGIGLILAGAY 91
: : : : : : : : : : : : : : : : : : : : : : : : :
Db 69 REGARMVLTGG---VLSALGNLSL-----LAFHD--GPSLLVGRFIVGLGV----- 110
: : : : : : : : : : : : : : : : : : : : : : : : :
QY 92 LDIKALDGLTKAAAFOWT---WKGVMLAGVVVTA-----VT-----RLTEILLPT 135
: : : : : : : : : : : : : : : : : : : : : : : : :
Db 111 -----GLVVAGTAGWAGRLGASGVTLGAILTAGFMWGPVITVSGLMGSTSIITPEA 163
: : : : : : : : : : : : : : : : : : : : : : : : :
QY 136 FANSYRNKLKNSINVALGFG-----EPSFDVAMGQSSALGFELSFKKS 178
: : : : : : : : : : : : : : : : : : : : : : : : :
164 IS-----VALSLIADVGVFGALGDARSTPS---ALGASSGIKHERSMKKA 204
: : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 3

```
US-09-815-242-13341
; Sequence 13341, Application US/09815242
; Patent No. US20020061569A1
```

GENERAL INFORMATION:

```
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELTRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
```

```
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13341
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13341
```

Query Match 8.9%; Score 80.5; DB 10; Length 650;

Best Local Similarity 25.7%; Pred. No. 2.2; Mismatches 45; Conservative 25; Mismatches 60; Indels 45; Gaps 10;

Matches 45; Conservative 25; Mismatches 60; Indels 45; Gaps 10;

```
QY 16 SSFAQAANDSKNGAF--GMSAGEKLLV-----YETSKQDPVIVPFLNLFGR---- 59
: : : : : : : : : : : : : : : : : : : : : : : : :
Db 401 STLAVSSGFLGALVCGFTAGALVLAIKKYVKVPRSEGAKSILLPLLTGTLTGFWMLA 460
: : : : : : : : : : : : : : : : : : : : : : : : :
QY 60 -----GIGSFAQGDILGSLILGFDVAGIGLILAGAYLDIKALD--GITKKAAFQWTW 112
: : : : : : : : : : : : : : : : : : : : : : : : :
Db 461 VNIPMAAINTAMNDFLGG--LGGSASVLLGIVLGG---MMAVDMGSPVYNKAAAY--VF 512
: : : : : : : : : : : : : : : : : : : : : : : : :
QY 113 GVMLA-----GVVTA-----VTRLTEILLPTTFANSYRNKLKNS--INVALG 153
: : : : : : : : : : : : : : : : : : : : : : : : :
Db 513 GTLATVSSGGSYMAAAMVAGMVPPLAIFVATLLFKDKFTKEERNSGLTNIIMG 567
: : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 4

```
US-09-738-626-4875
; Sequence 4875, Application US/09738626
; Publication No. US20020197605A1
```

GENERAL INFORMATION:

```
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4875
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4875
```

Query Match 8.9%; Score 80; DB 9; Length 487;

Best Local Similarity 22.3%; Pred. No. 1.7; Mismatches 37; Conservative 19; Mismatches 54; Indels 56; Gaps 6;

Matches 37; Conservative 19; Mismatches 54; Indels 56; Gaps 6;

```
QY 6 IFVATFCVFSFAQAANDSKNGAFGMSAGEKLLVYETSKQDPVIVPFLNLFGRFGIGSFA 65
: : : : : : : : : : : : : : : : : : : : : : : : :
```

```
Db 253 IMIVAFAFWF-----IYQSRKAE-----SGNDPLVPLEIFKRNFSLGNIC 297
QY 66 QGDILGSLILGFDVAGVIGLILAGAYLIDIKALDGTTRKKAQFQWTWKGVMVLAVVTVAVT 125
Db 298 -----IMAGFTVAGTTPIMLY-----FQAHGNNMARGFMVWVPOA 335
QY 126 RLTEIILPFTFANSYNRKLKNSLNLVAGLGGFEPFSDVAMGQSSALGF 171
Db 336 LMAAVLSPFV-----GKLVDRSN-----PGLMAALGF 362

RESULT 5
US-09-815-242-13924
; Sequence 13924, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13924
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Salmonella typhi

US-09-815-242-13924
Query Match 8.88; Score 79.5; DB 10; Length 319;
Best Local Similarity 27.58; Pred. No. 1.2;
Matches 39; Conservative 23; Mismatches 39; Indels 41; Gaps 9;

QY 39 VYETSKODPI-----VPFLNLFLG--FGIGSFAQGDILGSLILGFDVAGVIGLILAGA 90
Db 138 LFATYFQEPVGNILSNIPFVGALESFAGFGLIAGVIL-----ALMIPTYRAV 188
QY 91 YLDI-KALDGTTRKKAQFQWTWKGVMVLAVVTVAVTRLTEILPFTFANSYNRKLKNSLN 149
Db 189 MRDVFEGTPVMKESAY-----GI---GCTTWEV--IWRIVLPFT-----KNGV- 227
QY 150 VALGGFEPFSDVAMGQSSALGF 171
Db 228 --IGGIMLGLRALGETMAVTF 247

RESULT 6
US-09-738-626-6850
; Sequence 6850, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
```

```
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6850
; LENGTH: 417
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6850
Query Match 8.88; Score 79.5; DB 9; Length 417;
Best Local Similarity 24.78; Pred. No. 1.6;
Matches 39; Conservative 22; Mismatches 50; Indels 47; Gaps 8;

QY 47 PIVP-FLNLFLGFGIGSFAQGDILGSLILGFDVAGVIGLILAGA----YLDIKAL--DG 99
Db 177 PLLPLFIPIGIFLGMG-----MNGGLEIMSAFGVLLAVVGTLLFLAIQIILAGA 227
QY 100 ITKKAQFQWTWKGVMVLAVVTVAVTRLTEILPFTFANSYNRKLKNSLNLVAGLGGFEP-- 157
Db 228 VSKKN--PWKLFKN-MLPAYFTALGTSSAATIPVT---YQOTLKNDVDVNVAGFVVP 280
QY 158 -----SFDVAMGQSSALGFEL 173
Db 281 CATIHLAGSMKIGLTFVAVVFMVDMYGVGLSIGFL 318

RESULT 7
US-09-738-626-5811
; Sequence 5811, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5811
```

```
Query Match      8.7% Score 78.5; DB 9; Length 513;
Best Local Similarity 20.0%; Pred. No. 2.6;
Matches 38; Conservative 27; Mismatches 42; Indels 83; Gaps
9;
```

| | | | |
|----|-----|--|-----|
| QY | 49 | VP-FLLNFLGFGIGSFAQ--GDILGGSL--ILGFDAVGICLIILAGAYLD----- | 93 |
| Dd | | : :: : : : : : : : : | |
| Dd | 17 | VPAFLGITAVAGLGMGRSVGVVIGGAIKATGLFLLIGAGTTLVTASLEPLGAMINGAT | 76 |
| QY | 94 | -----IKALDGIITK---AAAFW----- | 108 |
| Dd | | : : : | |
| Dd | 77 | GMRGVPTNEAIGIAEYCAQVAWLMLGFAISLVLAFTNLRYYELTGHHVLFMSTM | 136 |
| QY | 109 | -----TW---KGWNLAGCVTMAYTRLTEIILPFTFANSYNRKLSNLVALGG | 154 |
| Dd | | : : : : : : : : : : | |
| Dd | 137 | LTIILATAGFDAMIWVGVGALLLGILMVS-----LP-AFAHPWTTRRTGGDDSIAGH | 187 |
| QY | 155 | FEPSEFDVAMG | 164 |
| Dd | | : : | |
| Dd | 188 | FGTAGYVAAG | 197 |

RESULT 9

```
Query Match      8.6%; Score 77.5; DB 9; Length 513;
Best Local Similarity 26.0%; Pred.No. 3.3;
Matches 32; Conservative 18; Mismatches 30; Indels 43; Gaps 5;
```

QY 3 KLIFVLATCVTSPPAQANDSRNCAFMSAGEKLLVYETSKODPIVPFLNLFLGFGIG 62
 :|||:::||| | | | | | | | | | | | |
Db 61 EILGIULSLRSGFAGATSOR-----IGRLTVV----- 92
 |||:::||| | | | | | | | | | | | |
QY 63 SFAQGDILGGLIPLGFDAVGIGLILAGAYLDIKALDGTKKAAFQWTGWKMLAGVVTM 122
 |||:::||| | | | | | | | | | | | |

Db 93 -FAAIVFFVGSLLMGF-AVNYGMMLMAGRFV---AGVG-----GYGGMIAPIVYTA 137

Qy 123 AVT 125

Db 138 EIS 140

RESULT 11

US-09-815-242-11951
Sequence 11951, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

PRIOR FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 11951

LENGTH: 570

TYPE: PRT

ORGANISM: pseudomonas aeruginosa

US-09-815-242-11951

Query Match 8.3%; Score 74.5; DB 10; Length 570;
Best Local Similarity 22.9%; Pred. No. 7.7;
Matches 39; Conservative 25; Mismatches 49; Indels 57; Gaps 8;

Qy 27 GAGMSAGEKLLVYETSKQDPIVPFLNLFLGFGIGSFAQ-----GDIL 70

Db 175 GAFAGVNLRLIVTG-----LHILNLNVFVFGFTDDETGRVVTGDLARFAGDPK 228

Qy 71 GGSILG-FDAVGIGL-----ILAGAYLDIKA---LDGITKKAFFOWT 109

Db 229 GGQFMAGNFPWMFGLPAACLAAYVNAAPRRKLIIGLLLSMALTAFLTGVTEPVEAFM 288

Qy 110 WKGV-----MLAGVVTMAVTRTEITLPTTFANSYNRKLNLSNVALG 153

Db 289 FLAPLLVLLHALLTG-LSMALTDLLDIRLGFTFSG-----GAIDLALG 330

RESULT 12

US-09-738-626-5024

Sequence 5024, Application US/09738626

Publication No. US20020197605A1

GENERAL INFORMATION:

APPLICANT: NAKAGAWA, SATOSHI

APPLICANT: MIZOGUCHI, HIROSHI

APPLICANT: ANDO, SEIKO

APPLICANT: HAYASHI, MIKIRO

APPLICANT: OCHIALI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-125

CURRENT APPLICATION NUMBER: US/09/738,626

CURRENT FILING DATE: 2000-12-18

PRIOR APPLICATION NUMBER: JP 99/377484

PRIOR FILING DATE: 1999-12-16

PRIOR APPLICATION NUMBER: JP 00/159162

PRIOR FILING DATE: 2000-04-07

PRIOR APPLICATION NUMBER: JP 00/280988

PRIOR FILING DATE: 2000-08-03

NUMBER OF SEQ ID NOS: 7059

SOFTWARE: PatentIn ver. 3.0

SEQ ID NO 5024

LENGTH: 708

TYPE: PRT

ORGANISM: Corynebacterium glutamicum

US-09-738-626-5024

Query Match 8.3%; Score 74.5; DB 9; Length 708;

Best Local Similarity 24.7%; Pred. No. 10;

Matches 38; Conservative 21; Mismatches 54; Indels 41; Gaps 8;

Qy 4 LLIFVLATFCVFSFAQANDSKNGAF-----MSAGEKLLVYETSKQDPIVPFL----- 52

Db 532 ILVFIAATIQVRRSF-----YGHKLWLQILVLGLTLTSLAISIVTTTTFIGRLIA 581

Qy 53 --LNLFLGFGIGSFAQG--DILGSLILG--FDA-----VGIGLILAG-----AYLDIK 95

Db 582 GNAGMMLGFLGHATVSVILMWVIAAALMLNKLKLDAPGALMTGVLAVAGTFLKLVFFDLV 641

Qy 96 ALDGITKKAFFOWTGWKGVMLAGVVTMAVTRUTE 129

Db 642 ALSGVPRAIAPLLS--GIALLTIAAMRGRTSE 672

RESULT 13

US-09-815-242-10165

Sequence 10165, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 20:13:37 ; Search time 33.8253 Seconds
(without alignments)
7815.311 Million cell updates/sec

Title: US-09-508-487-20
Perfect score: 862
Sequence: 1 gatttttcattgatccag.....atatattttctctgctaa 862

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| C 1 | 49.2 | 5.7 | 6243 | 2 | US-09-056-075-1 |
| 2 | 49 | 5.7 | 1422 | 1 | US-08-319-704-5 |
| 3 | 49 | 5.7 | 3077 | 3 | US-09-410-028-3 |
| 4 | 47.6 | 5.5 | 6243 | 2 | US-09-056-075-1 |
| 5 | 47.2 | 5.5 | 2251 | 4 | US-08-991-677-11 |
| 6 | 46.2 | 5.4 | 2960 | 3 | US-08-913-842-3 |
| 7 | 45.8 | 5.3 | 734 | 3 | US-09-014-583-1 |
| 8 | 45.8 | 5.3 | 1667 | 1 | US-08-485-284-1 |
| 9 | 45.2 | 5.2 | 782 | 4 | US-08-998-416-224 |
| C 10 | 45.2 | 5.2 | 3618 | 1 | US-07-872-678A-36 |
| C 11 | 44.2 | 5.1 | 20674 | 4 | US-09-641-638-651 |
| 12 | 44 | 5.1 | 837 | 4 | US-08-998-416-288 |
| 13 | 44 | 5.1 | 854 | 4 | US-08-998-416-534 |
| C 14 | 44 | 5.1 | 1850 | 3 | US-08-617-860B-32 |
| C 15 | 44 | 5.1 | 4098 | 2 | US-08-605-106-4 |
| 16 | 44 | 5.1 | 5064 | 4 | US-08-936-107A-8 |
| C 17 | 43.8 | 5.1 | 1972 | 1 | US-08-463-048-1 |
| C 18 | 43.8 | 5.1 | 1972 | 1 | US-08-463-229-1 |
| C 19 | 43.8 | 5.1 | 19124 | 2 | US-08-302-891-1 |
| C 20 | 43.8 | 5.1 | 19124 | 2 | US-08-487-826B-13 |
| 21 | 43.6 | 5.1 | 688 | 4 | US-08-998-416-972 |
| 22 | 43.6 | 5.1 | 4818 | 3 | US-08-817-926-27 |
| 23 | 43.4 | 5.0 | 246240 | 2 | US-08-724-394A-20 |
| 24 | 43.4 | 5.0 | 246240 | 2 | US-08-724-394A-21 |
| 25 | 43.4 | 5.0 | 246240 | 2 | US-08-724-394A-22 |
| C 26 | 42.8 | 5.0 | 98844 | 4 | US-09-791-211-10 |
| 27 | 42.4 | 4.9 | 658 | 4 | US-08-998-416-595 |

| | | | | | |
|------|------|-----|-------|---|--------------------|
| C 28 | 42.4 | 4.9 | 740 | 4 | US-08-998-416-971 |
| 29 | 42.4 | 4.9 | 3138 | 1 | US-07-867-106-4 |
| 30 | 42.4 | 4.9 | 5852 | 1 | US-07-867-106-2 |
| C 31 | 42 | 4.9 | 3395 | 4 | US-09-103-478-3 |
| C 32 | 42 | 4.9 | 3395 | 4 | US-09-103-478-3 |
| C 33 | 42 | 4.9 | 7560 | 4 | US-09-103-478-4 |
| C 34 | 42 | 4.9 | 7560 | 4 | US-09-103-478-4 |
| 35 | 41.8 | 4.8 | 636 | 4 | US-08-998-416-1137 |
| 36 | 41.8 | 4.8 | 1850 | 3 | US-08-617-860B-32 |
| 37 | 41.8 | 4.8 | 4098 | 2 | US-08-605-106-4 |
| C 38 | 41.6 | 4.8 | 1826 | 4 | US-09-286-691-11 |
| C 39 | 41.6 | 4.8 | 1826 | 4 | US-09-687-147-11 |
| 40 | 41.6 | 4.8 | 3618 | 1 | US-07-872-678A-36 |
| 41 | 41.4 | 4.8 | 8920 | 2 | US-08-446-855A-1 |
| 42 | 41.4 | 4.8 | 8920 | 4 | US-09-150-741-1 |
| C 43 | 41.4 | 4.8 | 11459 | 4 | US-09-462-136-7 |
| 44 | 41 | 4.8 | 663 | 4 | US-08-998-416-191 |
| 45 | 41 | 4.8 | 860 | 4 | US-08-998-416-287 |

ALIGNMENTS

RESULT 1
US-09-056-075-1/c
; Sequence 1, Application US/09056075
; Patent No. 5955368
; GENERAL INFORMATION:
; APPLICANT: Johnson, Eric A.
; APPLICANT: Bradshaw, Marite
; APPLICANT: Rood, Julian
; TITLE OF INVENTION: Expression System for Clostridium
; TITLE OF INVENTION: Species
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/056,075
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95238
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6243 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3770..4013
; OTHER INFORMATION: /note= "RP4 origin of DNA transfer (orit) from
; OTHER INFORMATION: plasmid RP4"
; US-09-056-075-1

Query Match 5.7%; Score 49.2; DB 2; Length 6243;
Best Local Similarity 58.0%; Pred. No. 0.046;

| | Query Match | 5.7% | Score 49; | DB 1; | Length 1422; |
|----|-----------------------|---|------------------|-----------|--------------|
| | Best Local Similarity | 46.8%; | Pred. No. 0.038; | | |
| | Matches 154; | Conservative 0; | Mismatches 175; | Indels 0; | Gaps 0; |
| QY | 56 | AAAAGAAATTCGAAATTCGTAATTTTATGGTCAAAATCAAGAAGCTCTATTGGSAAGCGAATT | 115 | | |
| Db | 91 | ATAATTATTATTGGATAAAATTTATAGCATAGATTATTATCTCTATATAGAAAAATAT | 150 | | |
| QY | 116 | TCAAGTAATACTTTGAAAAAGTTAAATTTAAATAGTTTAAAAACCTTTTAAATTC | 175 | | |
| Db | 151 | TAAAAAAATATTTTATAACCTTTTATAAAATTTTAAATTTTATTATTAAATTTTCTATCAA | 210 | | |
| QY | 176 | ATTAAATAGTTACTATAAATACCAAGTTTAAATAAGAGGGTTTTTATGAATAAAATTTTAAAT | 235 | | |

| Query Match | 5.7% | Score 49 | DB 3 | Length 3077 | |
|-----------------------|--------------|--|----------------|-------------|--------|
| Best Local Similarity | 48.4% | Pred. No. 0.044 | | | |
| Matches 136 | Conservative | 0 | Mismatches 145 | Indels 0 | Gaps 0 |
| QY | 95 | AGCTCTATTGGGAGCGCAATTCAAGTAATACTTTGAAAAAGTTAAATTTAAATAGTTT | 154 | | |
| Db | 1568 | AGTTAAATGAGACTGAGACACTTTTATGGAAGTTTAGTCTTAAGTTGATTTAGCTGTGCA | 1627 | | |
| QY | 155 | TAAAAACCTTTTTTAAATTTTCATTAATGATTTACTATAATACCGAGTTTTTAATAAGAGGT | 214 | | |
| Db | 1628 | GAGAAGCTCAATTTATTTTACTTTTCTATTTCATCTTATAAGTGTCTTATGAAAAAGT | 1687 | | |
| QY | 215 | TTTTTATCAATTAATTTTTTAAATTTGTTTGTCTAGCCCTTTTGTTTTTCTAGCTTTCG | 274 | | |
| Db | 1688 | TGATCTAAACAGGACTTAAATGTTTTTAACTAACCTTCTCTCTATGTTGTTTTTA | 1747 | | |
| QY | 275 | TCAAGCTGATGATTCTTAAAGCGCTTTTAAATTTGGGAGCGGAGAAAAACCTTTTAGCTTA | 334 | | |
| Db | 1748 | TGAACCTTAAAGTTTATCTAACATTGCTTAAATGATACAAGCTAGTGGAAAGTTGGGAATT | 1807 | | |
| QY | 335 | TGAAACTFAGTAAGAAAGATCCTATTGTGCCATTTTTATTGA | 375 | | |

Db 1808 AGAAATTTTGAACAGTTATACAGTTACAGGGTTGGTGGTCA 1848
||||| | | | | | | | | | | | | | | | | |

RESULT 4

US-09-056-075-1

; Sequence 1, Application US/09056075

; Patent No. 5955368

; GENERAL INFORMATION:

; APPLICANT: Johnson, Eric A.

; APPLICANT: Bradshaw, Marite

; APPLICANT: Rood, Julian

; TITLE OF INVENTION: Expression System for Clostridium

; TITLE OF INVENTION: Species

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Quarles & Brady

; STREET: 1 South Pluckney Street

; CITY: Madison

; STATE: WI

; COUNTRY: US

; ZIP: 53701-2113

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/056,075

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Seay, Nicholas J.

; REGISTRATION NUMBER: 27386

; REFERENCE/DOCKET NUMBER: 960296.95238

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 608-251-5000

; TELEFAX: 608-251-9166

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 6243 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: 3770..4013

; OTHER INFORMATION: /note= "RP4 origin of DNA transfer (orig)" from

; OTHER INFORMATION: plasmid RP4"

US-09-056-075-1

Query Match 5.5%; Score 47.6; DB 2; Length 6243;
Best Local Similarity 46.9%; Pred. No. 0.1;
Matches 183; Conservative 0; Mismatches 204; Indels 3; Gaps 1;

QY 19 AGAATTTGTAGATTTTCGACAATAAGACATTTAAAGAAATTCGAATTCGTAATTT 78
| | | | | | | | | | | | | | | | | |

Db 2733 AATTTATATCATATTTTCAATAATTTATTAATCTATTTTATGTTTTTAAATTTATTTATTTAA 2792
| | | | | | | | | | | | | | | | | |

QY 79 TATGGTCAAAATCAAGAGCTCTATTGGGAAGCGAATTTCAAGTAATACTTTGAAAAAAGT 138
| | | | | | | | | | | | | | | | | |

Db 2793 TAAGTCCCAATTCGAACATTAATCTATTTTATGTTTTTAAATTTATTTATTTATTTAA 2852
| | | | | | | | | | | | | | | | | |

QY 139 TAAATTTAAATAGTTTAAACCTTTTAAATTTTCAATTAATTAATGTTACTATAATACCA 198
| | | | | | | | | | | | | | | | | |

Db 2853 TATTTAACTAAATTAATGATTCCTTTTAAATTTTACTATTCATTCCTCAAT--A 2909
| | | | | | | | | | | | | | | | | |

QY 199 GTTTTAAAGAGGTTTATGAATAAATTTTAAATGTTGTTGCTAGCCCTTTTGTG 258
| | | | | | | | | | | | | | | | | |

Db 2910 TATTACTATAATTTATTTACAAAATTAATTTCTTCATTTGTAATATTAGATGATTTACTA 2969
| | | | | | | | | | | | | | | | | |

QY 259 TTTTCTTCTAGCTTTGCTCAAGCTGATGATTCCTAAAGCGCTTTTAAATTTGGAGCGGAG 318
| | | | | | | | | | | | | | | | | |

Db 2970 ATTTTACTTTTATATATATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3029
| | | | | | | | | | | | | | | | | |

QY 319 AAAAATTTTACCTTATGAAACTAGTAAGAAGATCTTATTTGCGCATTTTATTGTAACC 378
| | | | | | | | | | | | | | | | | |

Db 3030 TATAAATTTATGATTTATTTCCAAAGATATAAAGATTTTAATTTTCAATTTTAAACAATAC 3089
| | | | | | | | | | | | | | | | | |

QY 379 TTTTTTTAGGGTTTGGGAATAGGTTCTTTTG 408
| | | | | | | | | | | | | | | | | |

Db 3090 TTTTGTAAATATTAATGTTTAAATTTAATG 3119
| | | | | | | | | | | | | | | | | |

RESULT 5

US-08-991-677-11

; Sequence 11, Application US/08991677A

; Patent No. 6252135

; GENERAL INFORMATION:

; APPLICANT: Chiang, Vincent L

; APPLICANT: Carraway, Daniel T

; APPLICANT: Smeltzer, Richard H

; TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms

; FILE REFERENCE: 50617

; CURRENT APPLICATION NUMBER: US/08/991.677A

; CURRENT FILING DATE: 1997-12-16

; EARLIER APPLICATION NUMBER: US 60/033,381

; EARLIER FILING DATE: 1996-12-16

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 11

; LENGTH: 2251

; TYPE: DNA

; ORGANISM: Pinus taeda

US-08-991-677-11

Query Match 5.5%; Score 47.2; DB 4; Length 2251;
Best Local Similarity 49.6%; Pred. No. 0.1;
Matches 121; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 50 ATATTAAAGAAATTTGAATTTGCTAAATTTATGCTCAAAATCAAGAGCTCTATTGGGAAG 109
| | | | | | | | | | | | | | | | | |

Db 619 ACTATAAATTCAGTTGTGAATGAGTGTGTTATGAGATGTTTAAATTTATATA 678
| | | | | | | | | | | | | | | | | |

QY 110 CGAATTTCAAGTAATACTTTGAAAAAGTTAAATTTAAATAGTTTAAACCTTTTATA 169
| | | | | | | | | | | | | | | | | |

Db 679 TGTAATTAATTTATTTTGAATAACAATAATTAATTTGGATAAAAATGTTTGT 738
| | | | | | | | | | | | | | | | | |

QY 170 AATTTCAATTAATGTTTACTATATAATACCAGTTTAAATAAGAGAGGTTTATGAATAA 229
| | | | | | | | | | | | | | | | | |

Db 739 AATTTAGAGTAAATAATTTCAAAATCTAAAATAATTAACACACTATTATTTAAAAA 798
| | | | | | | | | | | | | | | | | |

QY 230 TTTAATTTGTTGTTGCTAGCCCTTTTGTGTTTTTTCTAGCTTTGCTCAAGCTGATGATC 289
| | | | | | | | | | | | | | | | | |

Db 799 TGTGTTGTAATTTTATCTATATTAAGTTAAAAATTTAGAAAAAATTAATTTAAATA 858
| | | | | | | | | | | | | | | | | |

QY 290 TAAA 293
| | | |

Db 859 TAAA 862
| | | |

RESULT 6

US-08-913-842-3/C

; Sequence 3, Application US/08913842

; Patent No. 6028250

; GENERAL INFORMATION:

; APPLICANT: OHBA, Toshiharu

; APPLICANT: TAKAHASHI, Shuichi

; APPLICANT: ANMA, Yoshiko

; APPLICANT: ASADA, Kiyozo

; APPLICANT: KATO, Ikunoshin

; TITLE OF INVENTION: PLANT PROMOTER AND METHOD FOR GENE

; TITLE OF INVENTION: EXPRESSION USING SAID PROMOTER

; NUMBER OF SEQUENCES: 75

; CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 7th Street N.W., Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913.842
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 07-073043
FILING DATE: 30-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/00777
FILING DATE: 26-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: OHBA-1
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2960 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-913-842-3

Query Match 5.4%; Score 46.2; DB 3; Length 2960;
Best Local Similarity 50.2%; Pred. No. 0.17;
Matches 114; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 30 AATTTTCGCAATTAAGACATTTATAAGAAATTCCTAATTTTGGTCAAT 89
DB 1732 AATTTTATATAATAAATCTAAAATCAATCAATATATATTTATGCAAGAG 1673
QY 90 CAAGAAGCTCTATTGGGAGCGAATTCAGTAATCTTTGAAAAAGTTAAAT 149
DB 1672 ATAGCAGGTAAATATCAAGCAGTTTAAATATATTTATGATCAATGTAATTTGTTAA 1613
QY 150 AGTTTAAAAACCTTTTAAATTTCAATTAATGTTACTATAATACCAAGTTTAAATAA 209
DB 1612 TTTCCTGTATTCAAATTTCTTATTATAATGTTGCTAATGATATAAAGGTTTAAACTAA 1553
QY 210 GAGGTTTTATCAATAAATTTTAAATGTTGTTGCTAGCCCTTTTG 256
DB 1552 GGTGTTTTAACTAAATTTGGGATATTGTTGGTCTTAAATCCATTTG 1506

RESULT 7
US-09-014-583-1
; Sequence 1, Application US/09014583
; Patent No. 6140077
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, Takeshi
; APPLICANT: SUZUKI, Tadashi
; APPLICANT: TOKUDA, Junko
; APPLICANT: KATO, No. 6140077uo
; APPLICANT: SAKAI, Yasuyoshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: TAKAHASHI, Hitoshi
; TITLE OF INVENTION: METHOD FOR PRODUCING PHYTASE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/014.583
FILING DATE: 28-JAN-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 200398/96
FILING DATE: 30-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 027066-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 734 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Candida boidinii
DEVELOPMENTAL STAGE: wild type
US-09-014-583-1

Query Match 5.3%; Score 45.8; DB 3; Length 734;
Best Local Similarity 49.8%; Pred. No. 0.16;
Matches 143; Conservative 0; Mismatches 142; Indels 2; Gaps 1;

QY 38 ACAATAAAGACATTTATTAAGAAATTTGAAATTCCTAATTTTATGTCATAACGAAGC 97
DB 425 ATAAAAAATTTATTTATTAATGAATTAATTTTCTTTTATTAATAATATCGTTAATCT 484
QY 98 TCTATTGGGACGCAATTTCAAGTAATCTTTGAAAAAGTTAAATTTAAATAGTTTAA 157
DB 485 TTTAAATCTTATTTATTTAATTTCTTTCTTTATCATAGTTATCATATACAA--TTATA 542
QY 158 AAACCTTTTAAATTTCAATTAATGTTACTATATATACCAAGTTTAAATAAGAGGTTT 217
DB 543 TAACATAGATACACAATTTATTTATTTATTTATTTATTTTAAATATTTTAAATATTG 602
QY 218 TATGAATAAATTTTAAATTTGTTGTTGCTAGCCTTTTGTGTTTCTAGCTTCTCA 277
DB 603 TTTAAATATATCTTAAATTTAATTTTACGAATATACAAATTTTAAAGACTTTCTTT 662
QY 278 AGCTGATGATCTTAAAGCGCTTTTAAATTTTGGGAGGGGAGAAAC 324
DB 663 TTTTACGAATTTTACGAACCTTTTAAAAAACAACAAAAAACAAC 709

RESULT 8
US-08-485-284A-1
; Sequence 1, Application US/08485284A
; Patent No. 5750372
; GENERAL INFORMATION:
; APPLICANT: SAKAI, YASUYOSHI
; APPLICANT: TANI, YOSHIKI
; APPLICANT: SHIBANO, YUJI
; APPLICANT: KONDO, HIROTO
; APPLICANT: HATANAKA, HARUYO
; TITLE OF INVENTION: NOVEL VECTOR HAVING PROMOTER THAT IS

TITLE OF INVENTION: INDUCIBLE BY METHANOL AND/OR GLYCEROL
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D. C.
COUNTRY: U.S.A.
ZIP: 20005-3918

APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Reibschung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/998,416

FILING DATE: 24-DEC-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: CH 0016/97

FILING DATE: 31-DEC-1996

ATTORNEY/AGENT INFORMATION:

NAME: Meigs, J. Timothy

REGISTRATION NUMBER: 38,241

REFERENCE/DOCKET NUMBER: PE/5-30306/A/CGC1976

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8587

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 224:

SEQUENCE CHARACTERISTICS:

LENGTH: 782 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:

ORGANISM: PAGL208UP

US-08-998-416-224

Query Match 5.2%; Score 45.2; DB 4; Length 782;

Best Local Similarity 48.1%; Pred. No. 0.22;

Matches 128; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

Qy 40 AATAAAGACATTTATAAAGAAATTCGAATTCCTAATTTATGGTCAATCAAGAGCTC 99

Db 464 AATAAAGAAATTTACTAGACTATAAATAAATAAATATATGAAGGAACATTTAATAATTA 405

Qy 100 TATTTGGGAGCGGAATTTCAAGTAATACTTTGAAAAAGTTAAATTTAAATAGTTTTTAAA 159

Db 404 TATATATCAGGGGAAGTAAATTAATAATTTATATAAATTAATATATATATATATATAT 345

Qy 160 ACCCTTTTAAATTCATTAATATGTTACTATAATACCAAGTTTAAATAAGAGGTTTTTA 219

Db 344 TTAATAAACTTAATAATCTATTTTATTAATAAATGGTATATATTAATTTAAATTAAT 285

Qy 220 TGAATAAATTTTAAATTTGTTGTTGCTAGCCCTTTTGTGTTGTTTCTAGCTTTGCTCAAG 279

Db 284 TAAATACATTTTAAATAAATTTCTATAGTAATTTCTTATTTTATTTTATAACATTTTAA 225

Qy 280 CTGATGATTTCTAAAGCGCTTTTAAAT 305

Db 224 ATGTTTATGTTAAATAGATAATAAT 199

RESULT 10

US-07-872-678A-36/C

; Sequence 36, Application US/07872678A

TITLE OF INVENTION: INDUCIBLE BY METHANOL AND/OR GLYCEROL
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:

ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.

STREET: 1100 New York Avenue, N.W.

CITY: Washington

STATE: D. C.

COUNTRY: U.S.A.

ZIP: 20005-3918

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485,284A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 43361/1992

FILING DATE: 28-FEB-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/025,416

FILING DATE: 01-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: WHITE JR., PAUL E.

REGISTRATION NUMBER: 32,011

REFERENCE/DOCKET NUMBER: 217755/FPS38209US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 861-3000

TELEFAX: (202) 822-0944

TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1667 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: genomic DNA

US-08-485-284A-1

Query Match 5.3%; Score 45.8; DB 1; Length 1667;

Best Local Similarity 49.8%; Pred. No. 0.19;

Matches 143; Conservative 0; Mismatches 142; Indels 2; Gaps 1;

Qy 38 ACAATAAAGACATTTATAAAGAAATTCGAATTCCTAATTTATGGTCAATCAAGAGC 97

Db 1366 ATAAAAAATTTATATTTAAATGAATTAATTCCTTTATTTTAAATAATATCGTTAATCT 1425

Qy 98 TCTATTGGGAAGCGGAATTTCAAGTAATACTTTGAAAAAGTTAAATTTAAATAGTTTTAA 157

Db 1426 TTTAAATCTATTTTATTTTAAATCTTTCTTTATCATCATAGTTATCATATAACAA--TTATA 1483

Qy 158 AAACCTTTTAAATTCATTAATATGTTACTATAATACCAAGTTTAAATAAGAGGTTTT 217

Db 1484 TAAATAGATACACAAATTTATTTATTTATTTATTTATTTTAAAAATTTGATTTATTT 1543

Qy 218 TATGAATAAATTTTAAATTTGTTGTTTCTAGCCCTTTTGTGTTTCTAGCTTTGCTCA 277

Db 1544 TTAATAATATCTTAAATTAATTTTATTTTACGAATATACAAATTTTAAACGACTTTCTTT 1603

Qy 278 AGCTGATGATTTCTAAAGCGCTTTTAAATTTGGGAGCGGAGAAAAAC 324

Db 1604 TTTTAAACGAATTTAAACGAATTTTAAAAAACAACAAAAAACAAC 1650

RESULT 9

US-08-998-416-224/C

; Sequence 224, Application US/08998416

; Patent No. 6239264

; GENERAL INFORMATION:

; APPLICANT: Philippsen, Peter

; APPLICANT: Pohlmann, Rainer

```
; Patent No. 5541060
; GENERAL INFORMATION:
; APPLICANT: Bell, Graeme, et al.
; TITLE OF INVENTION: DETECTION OF EARLY-ONSET
; TITLE OF INVENTION: NON-INSULIN-DEPENDENT DIABETES MELLITUS
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: Post Office Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/872,678A
; FILING DATE: 22-APRIL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coughlin, Daniel F.
; REGISTRATION NUMBER: 36,111
; REFERENCE/DOCKET NUMBER: ARCD016
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-787-1400
; TELEFAX: 713-789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3618 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-07-872-678A-36

Query Match 5.2%; Score 45.2; DB 1; Length 3618;
Best Local Similarity 52.1%; Pred. No. 0.29;
Matches 101; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 46 AGACATTATTAAGAATTGCAATGCTATTTTATGGTCAATCAAGAGCTCTATTGG 105
Db 3030 AGATGTTAATGAAGTTAATGCTTTAGTATAATAGACATAACAATAAGGATAAGTT 2971
106 GAAGCGAATTTCAAGTAATACCTTTGAAAAGTTAAATTTAAATAGTTTAAAAACCTTT 165
Db 2970 AAAGTTAAATTTAAATTAATGTTAAATTAACCTTTAAATTTTAAATATGTTAATAAT 2911
QY 166 TTTAAATTTCAATATATGTTACTATAATACCAAGTTTAAATGAAGGTTTATGTAATA 225
Db 2910 TTTTAAATTTAAATTAATTAATAATAATATGTTAAATTTAAATTAATGTTAAAT 2851
QY 226 AATTTTAAATGTT 239
Db 2850 AAACCTTAAATTTT 2837

RESULT 11
US-09-641-638-651
; Sequence 651, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueret, Ilyde
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BLALEDIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051Cp1
; CURRENT APPLICATION NUMBER: US/09/641,638

; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 651
; LENGTH: 20674
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1123..3123
; OTHER INFORMATION: 5'regulatory region
; NAME/KEY: exon
; LOCATION: 3124..3297
; OTHER INFORMATION: exon 1
; NAME/KEY: exon
; LOCATION: 3871..4072
; OTHER INFORMATION: exon 2
; NAME/KEY: exon
; LOCATION: 5552..5633
; OTHER INFORMATION: exon 3
; NAME/KEY: exon
; LOCATION: 5758..5880
; OTHER INFORMATION: exon 4
; NAME/KEY: exon
; LOCATION: 5996..6099
; OTHER INFORMATION: exon 5
; NAME/KEY: exon
; LOCATION: 6349..6509
; OTHER INFORMATION: exon 6
; NAME/KEY: exon
; LOCATION: 7379..7522
; OTHER INFORMATION: exon 7
; NAME/KEY: exon
; LOCATION: 8645..8854
; OTHER INFORMATION: exon 8
; NAME/KEY: exon
; LOCATION: 12254..12340
; OTHER INFORMATION: exon 9
; NAME/KEY: exon
; LOCATION: 12854..13023
; OTHER INFORMATION: exon 10
; NAME/KEY: exon
; LOCATION: 13308..13429
; OTHER INFORMATION: exon 11
; NAME/KEY: exon
; LOCATION: 16567..16667
; OTHER INFORMATION: exon 12
; NAME/KEY: exon
; LOCATION: 16775..16945
; OTHER INFORMATION: exon 13
; NAME/KEY: exon
; LOCATION: 17063..17554
; OTHER INFORMATION: exon 14
; NAME/KEY: misc_feature
; LOCATION: 17555..20674
; OTHER INFORMATION: 3'regulatory region
; NAME/KEY: allele
; LOCATION: 1128
; OTHER INFORMATION: 10-508-191 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 1182
; OTHER INFORMATION: 10-508-245 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 1559
; OTHER INFORMATION: 10-509-284 : polymorphic base C or T
```

```

RESULT 12
US-08-998-416-288
; Sequence 288, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippson, Peter
; APPLICANT: Pohmann, Rainer
; APPLICANT: Steiner, Sabine

```

```

; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 534:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 854 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1372UP
; US-08-998-416-534

Query Match          5.1%; Score 44; DB 4; Length 854;
Best Local Similarity 50.4%; Pred. No. 0.4;
Matches 134; Conservative 0; Mismatches 130; Indels 2; Gaps 1;

QY  40  AATAAAGACATTATTAAGAAGTAATGAAATGCTAAATTTATGGTCGAATCAAGAGCTC 99
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db   303  AATATAGAATATTACTAGACTAATAAATAAATAATATATGAAAGGTACATTTAATAATTA 362

QY  100  TATTGGGAAGGAATTTCAAGTATACCTTCAGAAAAAGTTAAATTTAAATAGTTTTAAAA 159
      |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db   363  TATATATCAATGAAGTAAATTAATAATTAATATATAAATAATTAATATATACCTTAATAT 422

QY  160  ACCTTTTTTAAATTCATTAATATGTTTACTATATAATACCAGTTTTTAATAAAGAGGTTTTTA 219
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db   423  TAATAAACTTAAT--AATCTATTTTAAATAAAATGGTATATTTAATATATAAANTTAA 480

QY  220  TGAATAAATTTTTAAATGTTGTTTTGCTAGCCTTTTGTGTTTTTTCTAGCTTGTCTCAAG 279

```

[illegible]

RESULT 14
US-08-617-850B-32/c
; Sequence 32, Application US/08617860B
; Patent No. 6133506
; GENERAL INFORMATION:
; APPLICANT: Typfer, R., Bautor, J., Bothmann, H., Filsak, E.,
; APPLICANT: Hvricks-Grandpierre, C., Klein, B., Martini, N.,
; APPLICANT: Mller, A., Schulte, W., Voetz, M., Walex, J.,
; APPLICANT: Schell, J.

| Query Match | 5.1%; | Score 44; | DB 3; | Length 1850; |
|-----------------------|-----------------|--|-----------|--------------|
| Best Local Similarity | 53.8%; | Pred. No. 0.46; | | |
| Matches 114; | Conservative 0; | Mismatches 95; | Indels 3; | Gaps |
| QY | 32 | TTTTCCACAAATAAGACACATATTATTAAGAAATTGCAATTGCTAAATTTATGTCAAATCA | 91 | |
| | | | | |
| Db | 414 | TTTTCATATTTCAAAAATTCAAAANAATATTATAAAACCTTTTAAATATATTTTAAACC | 355 | |
| QY | 92 | AGAAGCTCATTTGGGAAGCGAATTTCAAGTAATACATTGGAAGAAAGTTTAAATTTAAATAG | 151 | |
| | | | | |
| Db | 354 | AAAAAATATTTTAAAAATTTCAACAATATTTTAAAAATAAAAAAACTAAATTTTAAAAATA | 296 | |
| QY | 152 | TTTAAAAACCTTTTAAATTTTCATTAATATGTTACTATAATACACAGTTTAAATAAAGA | 211 | |
| | | | | |
| Db | 295 | --TTTAAAAAATTTTAAAAACCTAAATTTTATAATATTTTAAAAATATTTTTTTTAAAGAAT | 238 | |
| QY | 212 | GGTTTTATCAATAAAATTTTTTAAATGTCGTTT | 243 | |
| | | | | |
| Db | 237 | TTAAAAATATTTTAAATTTTAAAAAATATTTT | 206 | |

RESULT 15
 US-08-605-106-4/c
 : Sequence 4, Application US/08605106
 : Patent No. 5910631
 : GENERAL INFORMATION:
 : APPLICANT: Topfer, R.
 : APPLICANT: Martini, N.
 : APPLICANT: Schell, J.
 : TITLE OF INVENTION: MEDIUM CHAIN-SPECIFIC THIOESTERS
 : NUMBER OF SEQUENCES: 14
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
 : STREET: P.O. Box 2938
 : CITY: Minneapolis
 : STATE: MN
 : COUNTRY: USA
 : ZIP: 55402
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: DOS
 : SOFTWARE: FastSeq Version 2.0
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/605,106
 : FILING DATE: 23-SEPT-1996
 : CLASSIFICATION: 800
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: PCT/EP94/02935
 : FILING DATE: 01-MAR-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Woessner, Warren D
 : REGISTRATION NUMBER: 30,440
 : REFERENCE/DOCKET NUMBER: 235.001US1
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 612-373-6900
 : TELEFAX: 612-339-3061
 : TELEX:
 : INFORMATION FOR SEQ ID NO: 4:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 4098 Base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: double stranded
 : TOPOLOGY: linear
 : MOLECULE TYPE: : DNS (genomic)
 : HYPOTHETICAL: NO
 : ANTI-SENSE: NO
 : ORIGINAL SOURCE:
 : ORGANISM: Cuphea lanceolata
 : IMMEDIATE SOURCE:
 : LIBRARY: genomic Lambda FIX II
 : CLONE: C1Gt1
 : FEATURE:

```
NAME/KEY: CDS
LOCATION: join(1797..2294, 2658..2791, 2898..3011, 3132
FEATURE: ..3303, 3391..3459, 3672..3941)
NAME/KEY: Startcodon
LOCATION: 1797..1799
FEATURE:
NAME/KEY: exon II
LOCATION: 1787..2294
FEATURE:
NAME/KEY: intron II
LOCATION: 2295..2657
FEATURE:
NAME/KEY: exon III
LOCATION: 2658..2791
FEATURE:
NAME/KEY: intron III
LOCATION: 2792..2897
FEATURE:
NAME/KEY: exon IV
LOCATION: 2898..3011
FEATURE:
NAME/KEY: intron IV
LOCATION: 3012..3131
FEATURE:
NAME/KEY: exon V
LOCATION: 3132..3303
FEATURE:
NAME/KEY: intron V
LOCATION: 3304..3390
FEATURE:
NAME/KEY: exon VI
LOCATION: 3391..3459
FEATURE:
NAME/KEY: intron VI
LOCATION: 3460..3671
FEATURE:
NAME/KEY: exon VII
LOCATION: 3672..3941
FEATURE:
NAME/KEY: Stopcodon
LOCATION: 3942..3944
US-08-605-106-4
```

```
Query Match 5.1%; Score 44; DB 2; Length 4098;
Best Local Similarity 53.8%; Pred. No. 0.54;
Matches 114; Conservative 0; Mismatches 95; Indels 3; Gaps 1;

QY 32 TTTTCGACAAATAAGACATTTATTAAGAAGATTGAAATTCGTAATTTTATGCTCAATCA 91
Db 414 TTTTCAATATTTCAAAAATTCAAAAATATTTAAAACTTTTAAATATATTTTAAACC 355

QY 92 AGAAGCTCTATGGGAGCGAATTTCAAGTATCTTTGAAAAAGCTTAAATTTAAATAG 151
Db 354 AAAAAATATTTTAAAAATTCACAAATATTTTAAAAATAAAAAAACTAATTTTAAATA- 296

QY 152 TTTTAAAAACCTTTTAAATTCATTAATATGTTACTATATACCAGTTTAAATAAGA 211
Db 295 --TTTAAAAATTTTAAAAACCTAATTTTATATATTTTAAAAATATTTTAAAAAGAAT 238

QY 212 GCTTTTATGAATAAATTTTAAATGTTGTTT 243
Db 237 TTAATAATATTTAAATTTTAAAAAATATTTT 206
```

Search completed: March 15, 2003, 07:44:04
Job time : 71.8253 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 20:15:46 ; Search time 1448.7 Seconds
(without alignments)
12999.061 Million cell updates/sec

Title: US-09-508-487-22
Perfect score: 749
Sequence: 1 tagaattttcaacaaataaaa.....atttatctagaaaaatgggtg 749

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 24791104 seqs, 12571243825 residues 49582208
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US099D_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US101A_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US101B_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US102B_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result | No. | Score | Query Match % | Length | DB | ID | Description |
|--------|-----|-------|---------------|--------|----|----------------------|--------------------|
| | 1 | 749 | 100.0 | 749 | 15 | US-09-153-447-22 | Sequence 22, Appl |
| | 2 | 749 | 100.0 | 749 | 15 | US-09-156-447-22 | Sequence 22, Appl |
| | 3 | 749 | 100.0 | 749 | 19 | US-09-508-487-22 | Sequence 22, Appl |
| | 4 | 607 | 81.0 | 862 | 15 | US-09-153-447-20 | Sequence 20, Appl |
| | 5 | 607 | 81.0 | 862 | 15 | US-09-156-447-20 | Sequence 20, Appl |
| | 6 | 607 | 81.0 | 862 | 19 | US-09-508-487-20 | Sequence 18, Appl |
| | 7 | 558.2 | 74.5 | 759 | 15 | US-09-153-447-18 | Sequence 18, Appl |
| | 8 | 558.2 | 74.5 | 759 | 15 | US-09-156-447-18 | Sequence 18, Appl |
| | 9 | 558.2 | 74.5 | 759 | 19 | US-09-508-487-18 | Sequence 18, Appl |
| c | 10 | 98.4 | 13.1 | 1761 | 1 | PCT-US98-12764-63 | Sequence 63, Appl |
| c | 11 | 98.4 | 13.1 | 1761 | 32 | US-09-830-228-63 | Sequence 63, Appl |
| | 12 | 91.6 | 12.2 | 9399 | 1 | PCT-US98-12764-14 | Sequence 14, Appl |
| | 13 | 91.6 | 12.2 | 9399 | 32 | US-09-830-228-14 | Sequence 14, Appl |
| | 14 | 82.2 | 11.0 | 7074 | 1 | PCT-US98-12764-19 | Sequence 19, Appl |
| | 15 | 82.2 | 11.0 | 7074 | 32 | US-09-830-228-19 | Sequence 19, Appl |
| | 16 | 66 | 8.8 | 7110 | 42 | US-10-221-613-251 | Sequence 251, Appl |
| c | 17 | 61.4 | 8.2 | 417 | 17 | US-09-362-510-57521 | Sequence 57521, A |
| c | 18 | 61.4 | 8.2 | 417 | 17 | US-09-362-510A-57521 | Sequence 57521, A |
| c | 19 | 61.4 | 8.2 | 417 | 34 | US-09-904-013-57521 | Sequence 57521, A |
| c | 20 | 59.4 | 7.9 | 402 | 33 | US-09-873-402A-57873 | Sequence 57873, A |
| c | 21 | 59.4 | 7.9 | 402 | 64 | US-60-209-830-49195 | Sequence 49195, A |

58.8 7.9 588 31 US-09-821-837-2899 Sequence 2899, Ap
58 7.7 353 17 US-09-304-517A-267508 Sequence 267508,
58 7.7 353 17 US-09-371-146A-266370 Sequence 266370,
58 7.7 353 22 US-09-565-240-22020 Sequence 22020, A
58 7.7 353 25 US-09-654-617-71765 Sequence 71765, A
58 7.7 353 27 US-09-684-016-71765 Sequence 71765, A
58 7.7 353 37 US-09-985-678-267508 Sequence 267508,
58 7.6 567 33 US-09-873-402A-65625 Sequence 65625, A
57.2 7.6 472 25 US-09-654-617-422565 Sequence 422565,
57 7.6 472 27 US-09-684-016-422565 Sequence 422565,
57 7.6 472 63 US-60-197-872-1501 Sequence 1501, Ap
57 7.6 195257 66 US-60-226-176-2437 Sequence 2437, Ap
56.8 7.6 195257 67 US-60-233-468-2437 Sequence 2437, Ap
56.8 7.6 195257 75 US-60-313-371-2437 Sequence 53, Appl
56.6 7.6 9021 41 US-10-172-086-53 Sequence 681, App
56.4 7.5 547 27 US-09-698-014-681 Sequence 389, App
56.4 7.5 19380 42 US-10-221-613-389 Sequence 74654, A
56.2 7.5 633 33 US-09-874-708A-74654 Sequence 73521, A
56.2 7.5 633 65 US-60-211-750-73521 Sequence 190, App
55.8 7.4 112420 69 US-60-258-275-190 Sequence 51, Appl
55.4 7.4 14147 41 US-10-172-086-51 Sequence 5933, Ap
55 7.3 481 27 US-09-688-848-5933 Sequence 5933, Ap
55 7.3 481 59 US-60-159-589-5933 Sequence 40093, A
54.6 7.3 391 33 US-09-865-439A-40093

ALIGNMENTS

RESULT 1
US-09-153-447-22
; Sequence 22, Application US/09153447
; GENERAL INFORMATION:
; APPLICANT: BERGSTRÖM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3130.1
; CURRENT APPLICATION NUMBER: US/09/153,447
; EARLIER FILING DATE: 1998-09-15
; EARLIER FILING DATE: 1997-09-10
; EARLIER FILING DATE: 1997-09-10
; EARLIER FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 749
; TYPE: DNA
; ORGANISM: Borrelia garinii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (192)..(725)
US-09-153-447-22

Query Match 100.0%; Score 749; DB 15; Length 749;
Best Local Similarity 100.0%; Pred. No. 4.4e-122;
Matches 749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAGAATTTTCAACAAATAAGATATTGTTAAAGAAATTCGAAATTCGTAATTTATGTTA 60
DB 1 TAGAATTTTCAACAAATAAGATATTGTTAAAGAAATTCGAAATTCGTAATTTATGTTA 60
QY 61 AATCAAGAAGCTCTATTGGTAAAGCAATTCGAGTAACAAATTTGAAAAAGTTAAATTTA 120
DB 61 AATCAAGAAGCTCTATTGGTAAAGCAATTCGAGTAACAAATTTGAAAAAGTTAAATTTA 120
QY 121 AATAGTTCACAAAGCTTTTAAATTTTCAATTAATGCTACCATTAATACCAGTTTAATA 180
DB 121 AATAGTTCACAAAGCTTTTAAATTTTCAATTAATGCTACCATTAATACCAGTTTAATA 180
QY 181 AAGGGTTTTATGAATAAGTTTAAATTTTATTTTGTGTAATCTTTTGTGCTTTTCTA 240
DB 181 AAGGGTTTTATGAATAAGTTTAAATTTTATTTTGTGTAATCTTTTGTGCTTTTCTA 240

QY 241 GTTTTGTCTCAAGATGATCTTAAAGACACTTTTAACTGGGAGGGGAGAGAAAAATTTTGG 300
DB 241 GTTTTGTCTCAAGATGATCTTAAAGACACTTTTAACTGGGAGGGGAGAGAAAAATTTTGG 300
QY 301 TTTATGAACATAATAGAAAGATTCCTTGTACCATTTTATGAACCTTTTATTTAGGTT 360
DB 301 TTTATGAACATAATAGAAAGATTCCTTGTACCATTTTATGAACCTTTTATTTAGGTT 360
QY 361 TCGGGATAGGTTCTTTTGTCTCAAGGAGATATCTTGGAGGTTCTCTTATTTCTTGGATTG 420
DB 361 TCGGGATAGGTTCTTTTGTCTCAAGGAGATATCTTGGAGGTTCTCTTATTTCTTGGATTG 420
QY 421 ATCGGTTTGTATAGGTTTAACTACTTACAGAGCTTATTTGGACATCAAGGATTTTGATA 480
DB 421 ATCGGTTTGTATAGGTTTAACTACTTACAGAGCTTATTTGGACATCAAGGATTTTGATA 480
QY 481 ATAACTCTAAAGAGCTGATTTTAAAGTGGACTTGGGTAAGGAATGATTTGGCAGGTG 540
DB 481 ATAACTCTAAAGAGCTGATTTTAAAGTGGACTTGGGTAAGGAATGATTTGGCAGGTG 540
QY 541 TGGTTACTATGGCTGTGACAAAGATTGACAGAAATTTGTTCTTCCATTTACATTTGCTAATA 600
DB 541 TGGTTACTATGGCTGTGACAAAGATTGACAGAAATTTGTTCTTCCATTTACATTTGCTAATA 600
QY 601 ATTATACAGGAAGCTGAAATATAGTCTTAAATATAGCTTTGGAGGATTTGAGCCTAGTT 660
DB 601 ATTATACAGGAAGCTGAAATATAGTCTTAAATATAGCTTTGGAGGATTTGAGCCTAGTT 660
QY 661 TTGATATTACATGGGCAAGCTAGTGTCTTGTGTTTGGAGTGTCTTTTCAAAAAAGCT 720
DB 661 TTGATATTACATGGGCAAGCTAGTGTCTTGTGTTTGGAGTGTCTTTTCAAAAAAGCT 720
QY 721 ATTAATTTTATTTATCTAGAAAATGGGTG 749
DB 721 ATTAATTTTATTTATCTAGAAAATGGGTG 749

RESULT 2

US-09-156-447-22
; Sequence 22, Application US/09156447
; GENERAL INFORMATION:
; APPLICANT: BERGSTRÖM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3130.1
; CURRENT APPLICATION NUMBER: US/09/156,447
; CURRENT FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/059,036
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 1041/97
; PRIOR FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 749
; TYPE: DNA
; ORGANISM: Borrelia garinii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (192)..(725)
US-09-156-447-22

Query Match 100.0%; Score 749; DB 15; Length 749;
Best Local Similarity 100.0%; Pred. No. 4.4e-122;
Matches 749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAGAATTTTCAACAAATAAGATATTGTTAAAGAAATTCGAAATTCGTAATTTATGTTA 60
DB 1 TAGAATTTTCAACAAATAAGATATTGTTAAAGAAATTCGAAATTCGTAATTTATGTTA 60
QY 61 AATCAAGAAGCTCTATTGGTAAAGCAATTCGAGTAACAAATTTGAAAAAGTTAAATTTA 120
DB 61 AATCAAGAAGCTCTATTGGTAAAGCAATTCGAGTAACAAATTTGAAAAAGTTAAATTTA 120

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 22:48:52 ; Search time 48.9852 Seconds
(without alignments)
10731.744 Million cell updates/sec

Title: US-09-508-487-22

Perfect score: 749

Sequence: 1 tagaatttcaacaaataaa.....atttatctagaataatgggtg 749

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 501302 seqs, 350932545 residues

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|---------------------|
| C 1 | 52 | 6.9 | 416 | 10 | US-09-960-352-4584 |
| C 2 | 48.4 | 6.5 | 416 | 10 | US-09-960-352-4584 |
| C 3 | 48.2 | 6.4 | 2000 | 9 | US-09-938-842A-5004 |
| C 4 | 48 | 6.4 | 1044 | 9 | US-09-938-059-10 |
| C 5 | 48 | 6.4 | 1051 | 9 | US-09-938-059-22 |
| C 6 | 48 | 6.4 | 1140 | 9 | US-09-938-059-29 |
| C 7 | 47.2 | 6.3 | 640681 | 10 | US-09-790-988-1 |
| C 8 | 46.8 | 6.2 | 2251 | 10 | US-09-796-256A-11 |
| C 9 | 45.8 | 6.1 | 385 | 10 | US-09-960-352-1739 |
| C 10 | 45.8 | 6.1 | 419 | 10 | US-09-960-352-11234 |
| C 11 | 45.2 | 6.0 | 3272 | 10 | US-09-796-348-18 |
| C 12 | 44.2 | 5.9 | 376 | 10 | US-09-960-352-5087 |
| C 13 | 44.2 | 5.9 | 640681 | 10 | US-09-790-988-1 |
| C 14 | 44 | 5.9 | 513509 | 9 | US-09-754-853A-4 |
| C 15 | 43.8 | 5.8 | 516 | 10 | US-09-960-352-5785 |
| C 16 | 43.4 | 5.8 | 397 | 10 | US-09-960-352-13784 |
| C 17 | 43.4 | 5.8 | 411 | 10 | US-09-960-352-14521 |
| C 18 | 43.4 | 5.8 | 413 | 10 | US-09-960-352-2919 |
| C 19 | 43.4 | 5.8 | 431 | 10 | US-09-960-352-5558 |

| | | | | | |
|------|------|-----|--------|----|---------------------|
| C 20 | 43.4 | 5.8 | 335913 | 9 | US-09-754-853A-2 |
| C 21 | 43.4 | 5.8 | 335913 | 9 | US-09-754-853A-3 |
| C 22 | 43 | 5.7 | 479 | 10 | US-09-960-352-12872 |
| C 23 | 43 | 5.7 | 1367 | 9 | US-09-938-842A-4959 |
| C 24 | 43 | 5.7 | 4161 | 9 | US-10-098-841-135 |
| C 25 | 42.8 | 5.7 | 446 | 10 | US-09-960-352-3400 |
| C 26 | 42.8 | 5.7 | 335913 | 9 | US-09-754-853A-2 |
| C 27 | 42.8 | 5.7 | 335913 | 9 | US-09-754-853A-3 |
| C 28 | 42.6 | 5.7 | 449 | 10 | US-09-960-352-2412 |
| C 29 | 42.6 | 5.7 | 73308 | 10 | US-09-954-456-2276 |
| C 30 | 42.4 | 5.7 | 335 | 10 | US-09-960-352-13489 |
| C 31 | 42.4 | 5.7 | 203654 | 10 | US-09-820-905-3 |
| C 32 | 42 | 5.6 | 2000 | 9 | US-09-938-842A-3655 |
| C 33 | 42 | 5.6 | 2000 | 9 | US-09-938-842A-5073 |
| C 34 | 41.8 | 5.6 | 11474 | 10 | US-09-816-028A-1 |
| C 35 | 41.6 | 5.6 | 291 | 10 | US-09-960-352-1243 |
| C 36 | 41.6 | 5.6 | 317 | 10 | US-09-960-352-2699 |
| C 37 | 41.6 | 5.6 | 368 | 10 | US-09-834-975-47 |
| C 38 | 41.6 | 5.6 | 750 | 10 | US-09-770-149-9 |
| C 39 | 41.6 | 5.6 | 5046 | 10 | US-09-725-735A-13 |
| C 40 | 41.4 | 5.5 | 3303 | 10 | US-09-796-348-19 |
| C 41 | 41.4 | 5.5 | 3315 | 10 | US-09-796-348-17 |
| C 42 | 41.4 | 5.5 | 513509 | 9 | US-09-754-853A-4 |
| C 43 | 41.2 | 5.5 | 2251 | 10 | US-09-796-256A-11 |
| C 44 | 41 | 5.5 | 347 | 10 | US-09-960-352-3371 |
| C 45 | 41 | 5.5 | 1058 | 10 | US-09-804-682-30 |

ALIGNMENTS

RESULT 1

US-09-960-352-4584/C

; Sequence 4584, Application US/09960352

; Patent No. US20020137139A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Tao, Nengbing

; APPLICANT: Byatt, John C.

; APPLICANT: Mathialagan, Nagappan

; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AN

; FILE REFERENCE: 16511.006/37-21(10298)C

; CURRENT APPLICATION NUMBER: US/09/960.352

; CURRENT FILING DATE: 2001-09-24

; NUMBER OF SEQ ID NOS: 15112

; SEQ ID NO 4584

; LENGTH: 416

; TYPE: DNA

; ORGANISM: Bos taurus

; OTHER INFORMATION: Clone ID: 20-LIB3057-016-Q1-K1-E11

US-09-960-352-4584

Query Match 6.9%; Score 52; DB 10; Length 416;

Best Local Similarity 51.8%; Pred. No. 0.29;

Matches 118; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

Qy 95 TAACAAATTTGAAAAAGTTAAATTTAAATAGTCTCCAAAGCCTTTTAAATTCATTA 154

Db 288 TTAATAATTTTAAAAATTTTCTTTTAAAAATTTTAAAAATTTTAAAAATTTT 229

Qy 155 TATGCTACCAATAACACAGTTTAAATAAGGGTTTTTATGATAAGTTTATATTTAT 214

Db 228 TTTTATGAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 169

Qy 215 TTGTGTAATCTTTGCTGCTTTTCTAGTTTCTGCTCAAGATGATCTAAAGACCTTTAA 274

Db 168 TTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 109

Qy 275 TCTGGGACGGGAGAAAAATTTTGGTTTATGAACACTAATAAGAAAGA 322

Db 108 ATTTTTCATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 61

RESULT 2
US-09-960-352-4584
; Sequence 4584, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4584
; LENGTH: 416
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 20-LIB3057-016-Q1-K1-E11
US-09-960-352-4584

Query Match 6.5%; Score 48.4; DB 10; Length 416;
Best Local Similarity 51.4%; Pred. No. 1.3;
Matches 112; Conservative 0; Mismatches 106; Indels 0; Gaps 0;
QY 11 AACAAATAAGATATGTTAAAGAAATGAAATTCGTAATTTATGTTAAATCAAGAAG 70
DB 146 AATAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 205
QY 71 CTCATTGTAAGCAATTCGAGTACAAATTTGAAAGAAAGTTAAATTAATAGTTCCA 130
DB 206 AAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 265
QY 131 AAAGCCTTTTAAATTCATTAATGCTACCATATACCAGTTTAATAAAGGGGTTTT 190
DB 266 AGAAATTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAT 325
QY 191 TATGAATAAGTTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 228
DB 326 TAAATAAATAATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 363

RESULT 3
US-09-938-842A-5004
; Sequence 5004, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 5004
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-5004

Query Match 6.4%; Score 48.2; DB 9; Length 2000;
Best Local Similarity 51.7%; Pred. No. 1.8;

Matches 134; Conservative 0; Mismatches 123; Indels 2; Gaps 1;
QY 2 AGAATTTTCAACAAATAAAGATATTCGTTAAAGAAATGAAATTCGTAATTTATGTTAA 61
DB 186 AGTACTATCATTTAGTAAATCTCTCGAATATATAATATGATTCGAAATAATACATTA 245
QY 62 ATCAAGAAGCTCTATTGGTAAGCAATTCGAGTACAAATTTGAAAAAGTTTAAATTTAA 121
DB 246 ATTACCAATCTTTTCTCTATAATAAGTACCAGCTTAAATATACCATAGCAGAAGTTATA 305
QY 122 ATAGTTCCAAAAGCCCTTTTAAATTTCA--TTAATATGCTACCATTAATACCAAGTTTAA 179
DB 306 TTTCTAAACAGATCGTTTCTTAAATTTTCAATTTTAAATTTTAAATTTTAAATTTT 365
QY 180 AAAGGGGTTTTATGAATAAGTTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 239
DB 366 GTAATGGTATTTTAAATAATGATTATAAACTTGATATTTTCATATAGTAGTTTGTGTT 425
QY 240 AGTTTGGCTCAAGATGATT 258
DB 426 TTGTAGGATCATGTAAT 444

RESULT 4
US-09-998-059-10
; Sequence 10, Application US/09998059
; Publication No. US20030005485A1
; GENERAL INFORMATION:
; APPLICANT: Ohlrogge, John B.
; APPLICANT: Benning, Christoph
; APPLICANT: Gao, Hongbo
; APPLICANT: Girke, Thomas
; APPLICANT: White, Joseph A.
; TITLE OF INVENTION: Plant Seed Specific Promoters
; FILE REFERENCE: MSU-06689
; CURRENT APPLICATION NUMBER: US/09/998,059
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/250,401
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 1044
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-998-059-10

Query Match 6.4%; Score 48; DB 9; Length 1044;
Best Local Similarity 51.4%; Pred. No. 1.8;
Matches 111; Conservative 0; Mismatches 105; Indels 0; Gaps
QY 30 AAAAGAATTGAATTCGTAATTTTATGGTTAAATCAAGAAGCTCTATTGCTAAGCGAATT 89
DB 504 ATATGAATGCTATTTTGTCTTAAACCTTAAATTTGTACTATTGGAAGGAATTCATT 563
QY 90 TCGAGTAACAATTTGAAAAAGTTAAATTTAAATAGTTCACAAAGCCTTTTTTAAATTTTC 149
DB 564 TATTTTAACTGATGATTAGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTA 623
QY 150 ATTAATATGCTACCAATAACCAAGTTTAAATAAGGGGTTTTTATGAATAAGTTTAAAT 209
DB 624 TTAATTTCTTATTTATAAACAATAAAGTAATATTAATTTTCTTTTAAATAAATAATAT 683
QY 210 TTTATTTTGGTATCTTTTGTCTTTTCTAGTTTT 245
DB 684 TTTATTTTCAATAATTTATTTTGTCTTTTTTTTTTT 719

RESULT 5
US-09-998-059-22
; Sequence 22, Application US/09998059
; Publication No. US20030005485A1
; GENERAL INFORMATION:

```
; APPLICANT: Ohlrogge, John B.
; APPLICANT: Benning, Christoph
; APPLICANT: Gao, Hongbo
; APPLICANT: Girke, Thomas
; APPLICANT: White, Joseph A.
; TITLE OF INVENTION: Plant Seed Specific Promoters
; FILE REFERENCE: MSU-06689
; CURRENT APPLICATION NUMBER: US/09/998,059
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/250,401
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 1051
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-998-059-22

Query Match
Best Local Similarity 51.4%; Score 48; DB 9; Length 1051;
Matches 111; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

Qy 30 AAAAGAATTGAAATGCTTAATTTTATGCTTAAATCAAGAAGCTCTATTGGTAAGCGAATT 89
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 504 ATATGAATGCTATTTTCTCTTAAAACTTAAATTTGTAATTTTGAAGGAATTTCAATTT 563
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 90 TCGAGTAACAATTTGAAAAAGTTAAATTTAAATAGTTCGAAAGCCCTTTTAAATTTTC 149
    || || || || || || || || || || || || || || || || || || || ||
Db 564 TATTTTAAATGCTAGTAGATTTATATTAATTAATTAATTAATTAATTAATTAATTA 623
    || || || || || || || || || || || || || || || || || || || ||
Qy 150 ATTAATGCTACCATATACAGTTTAAATAAAGGGTTTTTATGAATAAGTTTTTAATT 209
    || || || || || || || || || || || || || || || || || || || ||
Db 624 TTAATCTTATTATAACAATAAGTAATTAATTTCTTAATTTAAATAACATAT 683
    || || || || || || || || || || || || || || || || || || || ||
Qy 210 TTTATTTGGTAATCTTTTGCTTTTCTTAGTTTT 245
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 684 TTTATTTTCATAAATTTATTTTGCTTTTTTTTTTTT 719
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
US-09-998-059-29
; Sequence 29, Application US/09998059
; Publication No. US20030005485A1
; GENERAL INFORMATION:
; APPLICANT: Ohlrogge, John B.
; APPLICANT: Benning, Christoph
; APPLICANT: Gao, Hongbo
; APPLICANT: Girke, Thomas
; APPLICANT: White, Joseph A.
; TITLE OF INVENTION: Plant Seed Specific Promoters
; FILE REFERENCE: MSU-06689
; CURRENT APPLICATION NUMBER: US/09/998,059
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/250,401
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 1140
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-998-059-29

Query Match
Best Local Similarity 51.4%; Score 48; DB 9; Length 1140;
Matches 111; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

Qy 30 AAAAGAATTGAAATGCTTAATTTTATGCTTAAATCAAGAAGCTCTATTGGTAAGCGAATT 89
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 557 ATATGAATGCTATTTTCTCTTAAAACTTAAATTTGTAATTTTGAAGGAATTTCAATTT 616
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 90 TCGAGTAACAATTTGAAAAAGTTAAATTTAAATAGTTCGAAAGCCCTTTTAAATTTTC 149
```

```
Db 617 TATTTTAAATGCTAGATTAGATTTTATAATTAATAATTAATAATGATGTAATAACTAAC 676
    || || || || || || || || || || || || || || || || || || || ||
Qy 150 ATTAATATGCTACCATATACAGTTTAAATAAAGGGTTTTTATGAATAAGTTTTTAATT 209
    || || || || || || || || || || || || || || || || || || || ||
Db 677 TTAATTTCTTATTTATAAACAATAAAGTAAATTTTCTTTAAATTAATAAATAACATAT 736
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 210 TTTATTTTGGTAATCTTTTGTGCTTTTCTAGTTTT 245
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 737 TTTATTTTCAATAATTTATTTGCTTTTTTTTTTTTT 772
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. US20020127687A1
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match
Best Local Similarity 51.5%; Score 47.2; DB 10; Length 640681;
Matches 134; Conservative 0; Mismatches 123; Indels 3; Gaps 1;

Qy 87 ATTCGAGTAACAATTTGAAAAAGTTAAATTTAAATAGTTCGAAAGCCCTTTTAAAT 146
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 546504 ATTTCAACAAAAATATATAAATAAATAATTTTATATAAATAAGTTCCTATTTTAA 546445
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 147 TTCATTAATATGCTACCATATACAGTTT---AATAAAGGGTTTTTATGAATAAGTTT 203
    || || || || || || || || || || || || || || || || || || || ||
Db 546444 ATAAATTTTGTAGAAATAATCACATTTTAAACATCTATTTTTCGCGAAAT 546395
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 204 TTAATTTTATTTTGGTAATCTTTTGTCTTTTCTAGTTTGTCTCAAGATGATCTAAA 263
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 546384 GTAAATTTTAAATATTAATATTTGATCGATGATGATTTGAAATTCCTATTTTAT 546325
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 264 AGCACTTTTAACTCGGAGCGGGAGAAAAATTTTGGTTTATGAACATAAAGAAGAT 323
    || || || || || || || || || || || || || || || || || || || ||
Db 546324 AGAAATAATTTCTCCTACAGAGCGGTATTTGTGTATAGGAATTAATCTCTGATGCC 546265
    || || || || || || || || || || || || || || || || || || || ||
Qy 324 TCTCTTGTCACCATTTTATTT 343
    || || || || || || || || || || || || || || || || || || || ||
Db 546264 ATACTTTTATATTTTATTT 546245
    || || || || || || || || || || || || || || || || || || || ||

RESULT 8
US-09-796-256A-11
; Sequence 11, Application US/09796256A
; Patent No. US20020078477A1
; GENERAL INFORMATION:
; APPLICANT: Chiang, Vincent L
; APPLICANT: Carraway, Daniel T
; APPLICANT: Smeltzer, Richard H
; TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms
; FILE REFERENCE: 50617/c-3532.0
; CURRENT APPLICATION NUMBER: US/09/796,256A
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US60/033381
```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:25:06 ; Search time 23 Seconds
(without alignments)
1006.876 Million cell updates/sec

Title: US-09-508-487-21
Perfect score: 904
Sequence: 1 MNKFLIVLLAFVCSFAQ.....DINMGQASALGFELFKKSY 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 598297 seqs, 130101922 residues

Total number of hits satisfying chosen parameters: 598297

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New.*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pcp.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pcp.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pcp.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pcp.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pcp.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pcp.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|----------------------|
| 1 | 89.5 | 9.9 | 450 | 6 | US-10-282-122A-69822 |
| 2 | 87 | 9.6 | 505 | 5 | US-09-134-000C-5340 |
| 3 | 87 | 9.6 | 505 | 5 | US-09-134-000C-5340 |
| 4 | 87 | 9.6 | 505 | 6 | US-10-282-122A-57091 |
| 5 | 83.5 | 9.2 | 442 | 6 | US-10-369-493-12272 |
| 6 | 83.5 | 9.2 | 650 | 6 | US-10-282-122A-73910 |
| 7 | 83 | 9.2 | 464 | 6 | US-10-282-122A-71751 |
| 8 | 81.5 | 9.0 | 692 | 6 | US-10-369-493-48 |
| 9 | 81.5 | 9.0 | 1291 | 6 | US-10-354-774-42 |
| 10 | 81 | 9.0 | 305 | 5 | US-09-988-067B-6 |
| 11 | 81 | 9.0 | 379 | 6 | US-10-156-761-11883 |
| 12 | 81 | 9.0 | 479 | 6 | US-10-282-122A-51378 |
| 13 | 79 | 8.7 | 486 | 6 | US-10-366-683-19571 |
| 14 | 78.5 | 8.7 | 372 | 6 | US-10-282-122A-78114 |
| 15 | 78.5 | 8.7 | 436 | 6 | US-10-282-122A-66257 |
| 16 | 78.5 | 8.7 | 663 | 6 | US-10-287-274-332 |
| 17 | 78.5 | 8.7 | 663 | 6 | US-10-282-122A-56452 |
| 18 | 78.5 | 8.7 | 663 | 6 | US-10-282-122A-75460 |
| 19 | 78.5 | 8.7 | 1291 | 6 | US-10-354-774-40 |
| 20 | 78 | 8.6 | 540 | 6 | US-10-156-761-13326 |
| 21 | 77 | 8.5 | 308 | 6 | US-10-282-122A-53673 |
| 22 | 77 | 8.5 | 447 | 6 | US-10-369-493-23305 |
| 23 | 77 | 8.5 | 792 | 6 | US-10-156-761-9442 |
| 24 | 76.5 | 8.5 | 571 | 1 | PCT-US02-32727-19854 |
| 25 | 76.5 | 8.5 | 571 | 5 | US-09-978-825-19854 |
| 26 | 76.5 | 8.5 | 571 | 6 | US-10-057-498-19854 |

Sequence 73132, A
Sequence 6859, Ap
Sequence 44444, A
Sequence 1078, Ap
Sequence 59036, A
Sequence 6417, Ap
Sequence 11538, A
Sequence 6426, Ap
Sequence 6426, Ap
Sequence 19270, A
Sequence 49274, A
Sequence 78316, A
Sequence 50724, A
Sequence 12622, A
Sequence 59641, A
Sequence 10352, A
Sequence 72971, A
Sequence 44555, A
Sequence 52786, A

ALIGNMENTS

RESULT 1
US-10-282-122A-69822
; Sequence 69822, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69822
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-10-282-122A-69822

Query Match 9.9%; Score 89.5; DB 6; Length 450;
Best Local Similarity 26.2%; Pred. No. 0.6;
Matches 49; Conservative 22; Mismatches 51; Indels 65; Gaps 11;
QY 5 LIVVLLAFC-VSSFAQADSKSA-----FNL-----GAG-- 33
Db 49 MVIAPITCTVVGAGTGMNMSKVGKTGYALLYFIVSTIALLLIGLIYVNVVQPGAGWN 108
QY 34 -----EKLAYETSKKD-PIVPLNLFGLFGIGSFAQGDILGGFLILGFDVAV-GIG 83
Db 109 IDVSTLDASKIAAYTAGDQDSIVGFIINVTIVGAFANGDIL--QVLMFSVIFGFA 165
QY 84 LILTGAY-----LD-----IKALDKNAPKAAFKWTWKGKMMLAGAV-TMAVTRL 126
Db 166 LHLRSGYGRPVLDIFDRFAHVFMNIIMIMIKLAPIGAF----GAMAFTIGAYGVSSLVQL 221
QY 127 TELIIPF 133
Db 222 GQLMICF 228
RESULT 2
US-09-134-000C-5340
; Sequence 5340, Application US/09134000C
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5340
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5340

Query Match 9.6%; Score 87; DB 5; Length 505;
Best Local Similarity 23.3%; Pred. No. 1.2;
Matches 45; Conservative 26; Mismatches 68; Indels 54; Gaps 10;
QY 6 IYVLLAFCVSSFAQA-----DDSKSAFNLGAGEKLLAYETSKKDPIVPLLN- 53
Db 85 IVALVIYALFPILQNTYTGLTIDPSLEAEAFGSKKEKLFKELOM---ALPFIISG 141
QY 54 -----LFLG-----FGIGSFAQGDILGGFLILGFDVAVGIGLITGAYLDIKALDKNA 100
Db 142 IRTATVLIIGTATLAALIGAGG-----LGTFFILLGIDRNNLSLIFIG-----ALSSAA 189
QY 101 PRAAFKW--TW-----GKGMMLAGAVTMAVTRLTEIIPFTFANSYNRKLKNSLNAFG- 152
Db 190 LAVLFNYGIHWLEKANGRLIIGGTI-----LGLLLGGFFWNQOTSSKEKOLTIAGKL 243
QY 153 GPEPSFDINMGOA 165
Db 244 GAEPDIINMYKA 256
RESULT 3
US-09-134-000C-5340
; Sequence 5340, Application US/09134000C
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778

; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5340
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5340
Query Match 9.6%; Score 87; DB 5; Length 505;
Best Local Similarity 23.3%; Pred. No. 1.2;
Matches 45; Conservative 26; Mismatches 68; Indels 54; Gaps 10;
QY 6 IYVLLAFCVSSFAQA-----DDSKSAFNLGAGEKLLAYETSKKDPIVPLLN- 53
Db 85 IVALVIYALFPILQNTYTGLTIDPSLEAEAFGSKKEKLFKELOM---ALPFIISG 141
QY 54 -----LFLG-----FGIGSFAQGDILGGFLILGFDVAVGIGLITGAYLDIKALDKNA 100
Db 142 IRTATVLIIGTATLAALIGAGG-----LGTFFILLGIDRNNLSLIFIG-----ALSSAA 189
QY 101 PRAAFKW--TW-----GKGMMLAGAVTMAVTRLTEIIPFTFANSYNRKLKNSLNAFG- 152
Db 190 LAVLFNYGIHWLEKANGRLIIGGTI-----LGLLLGGFFWNQOTSSKEKOLTIAGKL 243
QY 153 GPEPSFDINMGOA 165
Db 244 GAEPDIINMYKA 256
RESULT 4
US-10-282-122A-57091
; Sequence 57091, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1

RESULT 6
US-10-282-122A-73910

```
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78634
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71751
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Staphylococcus haemolyticus
; US-10-282-122A-71751

Query Match          9.2%; Score 83; DB 6; Length 464;
Best Local Similarity 23.9%; Pred. No. 2.8;
Matches 56; Conservative 31; Mismatches 77; Indels 70; Gaps 11;

QY 5 LTVLLAFVCSFSAQADD-----SKSAFNILGAGEKLLAYETSKKDPVFPFL--- 52
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 210 LVITLLSVNV--VTQAADRGFLSPMILSLGAVFLVAGILFIRELKTNPLVDFHIFKN 267
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 53 NLFGLFGIGSFAQGDILGGFLI-----LGFDAVIGGLI-----LTGAYLDIKALDKN 99
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 268 KGYSGATLSNFMNLGVAGGTIWNVTYQSLGFSLQAGLSITVLVAVLIMIRVGEKL 327
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 100 APKAAFK--WTWCKGMMLAGVMTAVTRTEI-----IIPF-----TFAN----- 137
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 328 LQKLGAKPLLVGSGFAFLGLLLLSLTLPLNVAYIITSSIGVLLFGIGLGTATPSTDFA 387
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 138 -----SYNRKLNKSNIAFG-----GFEPSP--DINMGQASALGF 170
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 388 VAEAPDDKVGAVSGLYKMASSIGNAFGVAMSTIYFGFAGFMNLQGGAGVLF 441
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8
US-10-369-493-48
; Sequence 48, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
```

```
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 48
; LENGTH: 692
; TYPE: PRT
; ORGANISM: Aquifex aeolicus
; US-10-369-493-48

Query Match          9.0%; Score 81.5; DB 6; Length 692;
Best Local Similarity 24.6%; Pred. No. 6.7;
Matches 44; Conservative 17; Mismatches 49; Indels 69; Gaps 10;

QY 39 YETSKKDPVFPFL-----LNFLFGIGSFAQGDILGGFLIIGFADAVIGLILTLTA 89
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 84 YEDTYKKPIV-FLSPYGLHHTLLIYAG-GLG-FLAGDILKSSDLGFLIGVGFYMQG 140
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 90 Y-----LDI-----KALDKNAPKAAFKW-----TWGK 111
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 141 YVKQIRVDGQEDLDAQNQKELMPVKVLDKEG-----KWLKCYVYVRDEKVFYGVW-- 193
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 112 GMMLAGVMTAVTRTEIIPFTFANSYRKLNKSNIAFGGPEPSFDINMGQASALGF 170
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 194 -----EVNVGKTKLYLLDTNVEENTPWREISSRL-----YVPDKDLRLRQQIIVLGF 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
US-10-354-774-42
; Sequence 42, Application US/10354774
; GENERAL INFORMATION:
; APPLICANT: Williams, James A.
; Thallev, Bruce S.
; TITLE OF INVENTION: Multivalent Vaccine For Clostridium
; Botulinum Neurotoxin
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/354,774
; FILING DATE: 30-Jan-2003
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/704,159
; FILING DATE: 28-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingollia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: OPND-02304
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1291 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 42:
US-10-354-774-42

Query Match          9.0%; Score 81.5; DB 6; Length 1291;
```

```
Best Local Similarity 23.8%; Pred. No. 15;
Matches 36; Conservative 24; Mismatches 56; Indels 35; Gaps 6;

Qy 47 IVPFLLNLFGLGSGFPAQGDILGGLILGFDVAV-----GIGLILGAYLD---- 92
| | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 620 IVPY-IGUALNVG-NETAKGNFENAFETAGASILLEFIPELLIPVVGAFLLSYIDNKNK 677
| | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Qy 93 -IKALDKNAPKAFAKWTGKGMLAGAVTMAVTRL-----TEIIIPFT 134
| | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 678 IIKIDNALTRNEKWSMDYGLIQAQNLSTVNTQPTTIKEGMYKALNQAOALEIILYR 737
| | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Qy 135 FANSYNRKLNLSNIAFGFPFSDINMGQA 165
| | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 738 Y-NIYSEKESKNINIDFINDSKNLNIGINQA 767
| | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :

RESULT 10
US-09-988-067B-6
; Sequence 6, Application US/09988067B
; GENERAL INFORMATION:
; APPLICANT: Haas, Rainer
; APPLICANT: Kleanthous, Harold
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Miller, Charles
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Odenbreit, Stefan
; APPLICANT: Meyer, Thomas
; TITLE OF INVENTION: Helicobacter Polypeptides and
; FILE OF INVENTION: Corresponding Polynucleotide Molecules
; FILE REFERENCE: 06132/040002
; CURRENT APPLICATION NUMBER: US/09/988.067B
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: US 08/831,309
; PRIOR FILING DATE: 1997-04-01
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-988-067B-6

Query Match 9.0%; Score 81; DB 5; Length 305;
Best Local Similarity 28.1%; Pred. No. 2.7;
Matches 57; Conservative 23; Mismatches 75; Indels 48; Gaps 14;

Qy 14 VFSSFAQADDSKAFN-LGAGEKLLAYETSKKDPVLPFLLNLFGLF---GIGSFAQGDIL 69
| | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 84 IASRFAGNGSGGLFNELSFYK---YFLGKK-RIIGFRHSILFFGYQLGGVGS-VPGSGL 138
| | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Qy 70 GGFLILGFD-----AVGIGLI---LTGAYLDIKALDKNAPKAFA 105
| | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 139 IVLPYGFNTDILLINWTKRASOKYVRRVKGUSIFYKDMTGRTGLDANTL-KRASRHVF 197
| | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Qy 106 KWTGK--GMLLAGAVTMAVTRLTEIIPFTFANSYN-RKLKNSLNIAFGGFPFSD--- 159
| | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 198 RKSSGLVIGMELGGSTWFSANLTL-----PFNQKSRITFQLQGRFGVWNDEYDIDRYG 253
| | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Qy 160 --INMGQASA-LGFEL-SFKKSY 178
| | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 254 DEIVLGGSSVELGVKVPAPKVN 276
| | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :

RESULT 11
US-10-156-761-11883
; Sequence 11883, Application US/10156761
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
```

```
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-282
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11883
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11883

Query Match 9.0%; Score 81; DB 6; Length 379;
Best Local Similarity 24.2%; Pred. No. 3.5;
Matches 37; Conservative 26; Mismatches 50; Indels 40; Gaps

Qy 27 AFNLGAGEKLL----AYETSKKDPVLPFLLNLFGLGIGSGFPAQGDILGGLILGFDVAVGI 82
| | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 240 ALALGLGEGVLGTSAY-TSKREQF-----GFPVATFQAVAV 275
| | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Qy 83 GLILTGAYLDIKALDKNAPKAFAKWTGKGMMLAGAVTMAVTRLTEIIPFTFANSYNRK 142
| | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 276 Q--AADRYIDLKAMEATLWQAARISTGAGGALPAAGDVAVAKI-----WASEGVRR 325
| | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Qy 143 LKNSLNIAFGGF--EPSFDINMGQASALGFELS 173
| | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 326 VVQTAQLHLGGFGADTDYPLHRYHAWAKHLELS 358
| | | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :

RESULT 12
US-10-282-122A-51378
; Sequence 51378, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
```



```

; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66257
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-66257

```

```

Query Match      8.7%; Score 78.5; DB 6; Length 436;
Best Local Similarity 37.1%; Pred. No. 7.4;
Matches 26; Conservative 7; Mismatches 24; Indels 13; Gaps 3;

QY  22 DDSKSAFNLGAGKLLAYETSKDPVLPFLNLFGLGCGSFAOGDILGGFLILGFDVAV- 80
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  114 DTSKIAAYAAAG-----KSTVDFLMNVIPGTVVGFANGDIL---QVLFFSVLF 161
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  81 GIGLILTCAY 90
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  162 GYALHRLGSY 171
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Search completed: March 14, 2003, 14:36:26
Job time : 25 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 20:13:37 ; Search time 29.3911 Seconds
(without alignments)
7815.311 Million cell updates/sec

Title: US-09-508-487-22

Perfect score: 749

Sequence: 1 tagaattttcaacaataaa.....atttatctagaagaatgggtg 749

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 51.4 | 6.9 | 20674 | 4 | US-09-641-638-651 |
| 2 | 49.6 | 6.6 | 665 | 2 | US-08-883-795A-36 |
| C 3 | 48.6 | 6.5 | 837 | 4 | US-08-998-416-288 |
| C 4 | 47.2 | 6.3 | 665 | 2 | US-08-883-795A-36 |
| C 5 | 46.8 | 6.2 | 2251 | 4 | US-08-991-677-11 |
| C 6 | 45.2 | 6.0 | 19124 | 2 | US-08-487-826B-13 |
| C 7 | 44.4 | 5.9 | 8920 | 2 | US-08-446-855A-1 |
| C 8 | 44.4 | 5.9 | 8920 | 4 | US-09-150-741-1 |
| C 9 | 44.2 | 5.9 | 5852 | 1 | US-07-867-106-2 |
| 10 | 42.8 | 5.7 | 615 | 4 | US-08-998-416-186 |
| 11 | 42.8 | 5.7 | 636 | 4 | US-08-998-416-1137 |
| 12 | 42.8 | 5.7 | 837 | 4 | US-08-998-416-288 |
| 13 | 42.8 | 5.7 | 1850 | 3 | US-08-617-860B-32 |
| C 14 | 42.8 | 5.7 | 3440 | 1 | US-08-471-791-27 |
| C 15 | 42.8 | 5.7 | 3440 | 5 | PCT-US91-01746-27 |
| C 16 | 42.8 | 5.7 | 4098 | 2 | US-08-605-106-4 |
| C 17 | 42.6 | 5.7 | 1850 | 3 | US-08-617-860B-32 |
| C 18 | 42.6 | 5.7 | 4098 | 2 | US-08-605-106-4 |
| C 19 | 42 | 5.6 | 731 | 1 | US-08-451-405A-2 |
| C 20 | 42 | 5.6 | 834 | 4 | US-08-998-416-305 |
| C 21 | 41.4 | 5.5 | 2960 | 3 | US-08-913-842-3 |
| C 22 | 41.2 | 5.5 | 2251 | 4 | US-08-991-677-11 |
| C 23 | 41 | 5.5 | 2663 | 1 | US-08-136-743B-3 |
| C 24 | 40.6 | 5.4 | 767 | 4 | US-08-998-416-472 |
| C 25 | 40.4 | 5.4 | 782 | 4 | US-08-998-416-224 |
| C 26 | 40.4 | 5.4 | 1454 | 2 | US-08-713-000-7 |
| C 27 | 40.4 | 5.4 | 1454 | 2 | US-08-975-316-7 |

C 28 40.4 5.4 1454 4 US-09-211-710-7 Sequence 7, Appli
C 29 40.4 5.4 1454 4 US-09-615-192A-7 Sequence 7, Appli
C 30 40.4 5.4 1460 4 US-09-615-192A-95 Sequence 95, Appli
C 31 40.4 5.4 1474 2 US-08-975-316-71 Sequence 71, Appli
C 32 40.4 5.4 1474 4 US-09-615-192A-71 Sequence 71, Appli
C 33 40.4 5.4 1474 4 US-09-615-192A-402 Sequence 402, App
C 34 40.4 5.4 1474 4 US-09-325-932A-202 Sequence 202, App
C 35 40.4 5.4 5303 4 US-08-971-395-4 Sequence 4, Appli
C 36 40.4 5.4 6124 4 US-08-213-419B-3 Sequence 3, Appli
C 37 40.4 5.4 6124 4 US-08-213-419B-3 Sequence 3, Appli
C 38 40.4 5.4 6243 2 US-09-056-075-1 Sequence 3, Appli
C 39 40 5.3 1422 1 US-08-319-704-5 Sequence 5, Appli
C 40 5.3 6243 2 US-09-056-075-1 Sequence 1, Appli
C 41 39.8 5.3 600 4 US-09-134-001C-1770 Sequence 1770, Ap
C 42 39.8 5.3 636 4 US-08-998-416-1137 Sequence 1137, Ap
C 43 39.8 5.3 2672 1 US-08-703-947-1 Sequence 1, Appli
C 44 39.8 5.3 3850 3 US-08-329-799-34 Sequence 34, Appli
C 45 39.6 5.3 662 4 US-08-998-416-185 Sequence 185, App

ALIGNMENTS

RESULT 1

US-09-641-638-651
; Sequence 651, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET.051CPI
; CURRENT APPLICATION NUMBER: US/09/641.638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 651
; LENGTH: 20674
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1123..3123
; OTHER INFORMATION: 5'regulatory region
; NAME/KEY: exon
; LOCATION: 3124..3297
; OTHER INFORMATION: exon 1
; NAME/KEY: exon
; LOCATION: 3871..4072
; OTHER INFORMATION: exon 2
; NAME/KEY: exon
; LOCATION: 5552..5633
; OTHER INFORMATION: exon 3
; NAME/KEY: exon
; LOCATION: 5758..5880
; OTHER INFORMATION: exon 4
; NAME/KEY: exon
; LOCATION: 5996..6099
; OTHER INFORMATION: exon 5
; NAME/KEY: exon
; LOCATION: 6349..6509
; OTHER INFORMATION: exon 6

```
NAME/KEY: exon
LOCATION: 7379..7522
OTHER INFORMATION: exon 7
NAME/KEY: exon
LOCATION: 8645..8854
OTHER INFORMATION: exon 8
NAME/KEY: exon
LOCATION: 12254..12340
OTHER INFORMATION: exon 9
NAME/KEY: exon
LOCATION: 12854..13023
OTHER INFORMATION: exon 10
NAME/KEY: exon
LOCATION: 13308..13429
OTHER INFORMATION: exon 11
NAME/KEY: exon
LOCATION: 16567..16667
OTHER INFORMATION: exon 12
NAME/KEY: exon
LOCATION: 16775..16945
OTHER INFORMATION: exon 13
NAME/KEY: exon
LOCATION: 17063..17554
OTHER INFORMATION: exon 14
NAME/KEY: misc_feature
LOCATION: 17555..20674
OTHER INFORMATION: 3'regulatory region
NAME/KEY: allele
LOCATION: 1128
OTHER INFORMATION: 10-508-191 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 1182
OTHER INFORMATION: 10-508-245 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 1559
OTHER INFORMATION: 10-509-284 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 1570
OTHER INFORMATION: 10-509-295 : deletion of C
NAME/KEY: allele
LOCATION: 1827
OTHER INFORMATION: 10-510-173 : variable motif ATTGA or TTTTIT
NAME/KEY: allele
LOCATION: 2048
OTHER INFORMATION: 10-511-62 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 2323
OTHER INFORMATION: 10-511-337 : insertion of T
NAME/KEY: allele
LOCATION: 2341
OTHER INFORMATION: 10-512-36 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 2623
OTHER INFORMATION: 10-512-318 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 2832
OTHER INFORMATION: 10-513-250 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 2844
OTHER INFORMATION: 10-513-262 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 2934
OTHER INFORMATION: 10-513-352 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 2947
OTHER INFORMATION: 10-513-365 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 3802
OTHER INFORMATION: 12-206-81 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 4062
OTHER INFORMATION: 10-343-231 : deletion of C
NAME/KEY: allele
LOCATION: 4088
OTHER INFORMATION: 12-206-366 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 4109
OTHER INFORMATION: 10-343-278 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 4170
OTHER INFORMATION: 10-343-339 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 5903
OTHER INFORMATION: 10-346-23 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6019
OTHER INFORMATION: 10-346-141 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6141
OTHER INFORMATION: 10-346-263 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 6183
OTHER INFORMATION: 10-346-305 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 6338
OTHER INFORMATION: 10-347-74 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6375
OTHER INFORMATION: 10-347-111 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 6429
OTHER INFORMATION: 10-347-165 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 6467
OTHER INFORMATION: 10-347-203 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6484
OTHER INFORMATION: 10-347-220 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6534
OTHER INFORMATION: 10-347-271 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 6611
OTHER INFORMATION: 10-347-348 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 7668
OTHER INFORMATION: 10-348-391 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 8608
OTHER INFORMATION: 10-349-47 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 8658
OTHER INFORMATION: 10-349-97 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 8703
OTHER INFORMATION: 10-349-142 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 8777
OTHER INFORMATION: 10-349-216 : deletion of CTG
NAME/KEY: allele
LOCATION: 8785
OTHER INFORMATION: 10-349-224 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 8926
OTHER INFORMATION: 10-349-368 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 12171
OTHER INFORMATION: 10-350-72 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 12429
OTHER INFORMATION: 10-350-332 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 13341
OTHER INFORMATION: 10-507-170 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 13492
```


FEATURE: ;
NAME/KEY: CAAT-Signal ;
LOCATION: 1428..1432 ;
FEATURE: ;
NAME/KEY: TATA-Signal ;
LOCATION: 1553..1556 ;
FEATURE: ;
NAME/KEY: Transcription start ;
LOCATION: 1585 ;
FEATURE: ;
NAME/KEY: Leguminbox ;
LOCATION: 1642..1657 ;
FEATURE: ;
NAME/KEY: Startcodon ;
LOCATION: 1797..1799 ;
FEATURE: ;
NAME/KEY: CDS ;
LOCATION: 1797..1850 ;
US-08-617-860B-32 ;
Query Match 5.7%; Score 42.8; DB 3; Length 1850;
Best Local Similarity 48.9%; Pred. No. 0.39;
Matches 114; Conservative 0; Mismatches 119; Indels 0; Gaps 0;
Qy 111 GTTAAATTAATAGTTCACAAAGCCCTTTTAAATTCATTAATATGCTACCAATAATAC 170
Db 166 GTAATATTTATGAATATTCGAATATTTTGAATTTTAAATATTTTAAATTTTAAATTTA 225
Qy 171 CAGTTTAATAAGGGTTTATGAATAAGTTTATTAATTTTATTTTATTTTGGTAACTTTTGT 230
Db 226 AATATATTTAAATCTTTTAAATAAATATTTTAAATATTTAAATATTTAGTTTAA 285
Qy 231 GCTTTTCTAGTTTGTCTCAAGATGATCTTAAAGCACTTTTAACTCTGGAGCGGAGAA 290
Db 286 ATTTTAAATATTTTAAATATTTTATTTTATTTTAAATATTTGTTGAATTTTAA 345
Qy 291 AAATTTTGGTTTATGAACCTAATAAGAAAGATCTCTGTGACCATTTTATTTAT 343
Db 346 ATATTTTGTCTTTTAAATATATTTTAAAGTTTAAATATTTTGAATTT 398
RESULT 14
US-08-471-791-27/c
; Sequence 27, Application US/08471791
; Patent No. 5723595
; GENERAL INFORMATION:
; APPLICANT: Thompson, Gregory A
; APPLICANT: Knauf, Vic C
; TITLE OF INVENTION: Plant Desaturases-Compositions
; TITLE OF INVENTION: and Uses
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: California
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.1
; SOFTWARE: Microsoft Word 5.1 (a)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,791
; FILING DATE: 6-JUNE-95
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/762,762
; FILING DATE: 16-SEPT-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/01746

; FILING DATE: 14-MAR-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/615,784
; FILING DATE: 14-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/567,373
; FILING DATE: 13-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/494,106
; FILING DATE: 16-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Lassen, Elizabeth
; REGISTRATION NUMBER: 31,845
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 69-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510
; TELEX: 350370 CGNE
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3440 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
US-08-471-791-27
Query Match 5.7%; Score 42.8; DB 1; Length 3440;
Best Local Similarity 52.1%; Pred. No. 0.45; Indels 8; Gaps 1;
Matches 125; Conservative 0; Mismatches 107;
Qy 1 TAGAATTTTCAACAAATAAAGATATTTGTTAAAGAAATTCGAATTCCTAATTTATGGTTA 60
Db 2405 TAAATTTTATAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2346
Qy 61 AATCAAGAAGCTCTATTTGGTAGCGAAATTCGAGTAACAATTTGAAAAAGTTAAATTA 120
Db 2345 GTTTCAAAAAATTTCAAAAAAATAAATTTCAAAAAAATAAATTTCAAAAAAAGGAATTTAA 2286
Qy 121 AATAGTTCACAAAGCCCTTTTAAATTTCAATTAATATGCTACCAATAATACCAGTTTAATA 180
Db 2285 AAAATTTTCGAA-----TCTGAAACATATAATCTGAACTATAAAAAAATTTCTTTT 223
Qy 181 AAGGGTTTTATGAATGAATTTTAAATTTTATTTTATTTGTTAATCTTTTGTGTTTTCTTA 240
Db 2233 TTTCATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTA 2174
RESULT 15
PCT-US91-01746-27/c
; Sequence 27, Application PC/TUS9101746
; GENERAL INFORMATION:
; APPLICANT: Thompson, Gregory A
; APPLICANT: Knauf, Vic C
; TITLE OF INVENTION: Plant Desaturases-Compositions and Uses
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: California
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 6.0.7
; SOFTWARE: Microsoft Word 4.0

Search completed: March 15, 2003, 07:44:33
Job time : 58.3911 secs

| Result No. | Score | Query | | Length | DB | ID | Description |
|------------|-------|-------|-----|--------|------------------|----|-------------------|
| | | Match | % | | | | |
| 1 | 904 | 100.0 | 178 | 15 | US-09-153-447-21 | | Sequence 21, Appl |
| 2 | 904 | 100.0 | 178 | 15 | US-09-156-447-21 | | Sequence 21, Appl |
| 3 | 904 | 100.0 | 178 | 19 | US-09-508-487-21 | | Sequence 21, Appl |
| 4 | 786.5 | 87.0 | 177 | 15 | US-09-153-447-23 | | Sequence 23, Appl |
| 5 | 786.5 | 87.0 | 177 | 15 | US-09-156-447-23 | | Sequence 23, Appl |
| 6 | 785.5 | 87.0 | 177 | 19 | US-09-508-487-23 | | Sequence 23, Appl |

QY 61 GSAQGDILGGFLILGFDVAVGIGLILTGAYLDIKALDKNAPKAAFKWTWCKGMMLAGAVT 120
|||||
Db 61 GSAQGDILGGFLILGFDVAVGIGLILTGAYLDIKALDKNAPKAAFKWTWCKGMMLAGAVT 120
|||||
QY 121 MAVTRLTEIIIPFTFANSYNRKLNLSNIAFGGFEPSFDINMQASALGFELSFKKSY 178
|||||
Db 121 MAVTRLTEIIIPFTFANSYNRKLNLSNIAFGGFEPSFDINMQASALGFELSFKKSY 178
|||||

RESULT 2

US-09-156-447-21

; Sequence 21, Application US/09156447

; GENERAL INFORMATION:

; APPLICANT: BERGSTROM, Sven

; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA

; FILE REFERENCE: 454312-3130.1

; CURRENT APPLICATION NUMBER: US/09/156,447

; CURRENT FILING DATE: 1998-09-15

; PRIOR APPLICATION NUMBER: 60/059,036

; PRIOR FILING DATE: 1997-09-10

; PRIOR APPLICATION NUMBER: 1041/97

; PRIOR FILING DATE: 1997-09-10

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 21

; LENGTH: 178

; TYPE: PRT

; ORGANISM: Borrelia afzelii

US-09-156-447-21

Query Match

Best Local Similarity 100.0%; Score 904; DB 15; Length 178;

Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKFLIVLLAFVFSFAQADDSKSAFNLAGGKLLAYETSKKDPVLPVLLNLFGLFGI 60
|||||
Db 1 MNKFLIVLLAFVFSFAQADDSKSAFNLAGGKLLAYETSKKDPVLPVLLNLFGLFGI 60
|||||
QY 61 GSAQGDILGGFLILGFDVAVGIGLILTGAYLDIKALDKNAPKAAFKWTWCKGMMLAGAVT 120
|||||
Db 61 GSAQGDILGGFLILGFDVAVGIGLILTGAYLDIKALDKNAPKAAFKWTWCKGMMLAGAVT 120
|||||

QY 121 MAVTRLTEIIIPFTFANSYNRKLNLSNIAFGGFEPSFDINMQASALGFELSFKKSY 178
|||||

Db 121 MAVTRLTEIIIPFTFANSYNRKLNLSNIAFGGFEPSFDINMQASALGFELSFKKSY 178
|||||

RESULT 3

US-09-508-487-21

; Sequence 21, Application US/09508487

; GENERAL INFORMATION:

; APPLICANT: BERGSTROM, Sven

; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA

; FILE REFERENCE: 454312-3150

; CURRENT APPLICATION NUMBER: US/09/508,487

; CURRENT FILING DATE: 2000-03-10

; PRIOR APPLICATION NUMBER: 60/059,036

; PRIOR FILING DATE: 1997-09-10

; PRIOR APPLICATION NUMBER: 1041/97

; PRIOR FILING DATE: 1997-09-10

; PRIOR APPLICATION NUMBER: PCT/IB98/01424

; PRIOR FILING DATE: 1998-09-04

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 21

; LENGTH: 178

; TYPE: PRT

; ORGANISM: Borrelia afzelii

US-09-508-487-21

Query Match

Best Local Similarity 100.0%; Score 904; DB 19; Length 178;

Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKFLIVLLAFVFSFAQADDSKSAFNLAGGKLLAYETSKKDPVLPVLLNLFGLFGI 60
|||||
Db 1 MNKFLIVLLAFVFSFAQADDSKSAFNLAGGKLLAYETSKKDPVLPVLLNLFGLFGI 60
|||||
QY 61 GSAQGDILGGFLILGFDVAVGIGLILTGAYLDIKALDKNAPKAAFKWTWCKGMMLAGAVT 120
|||||
Db 61 GSAQGDILGGFLILGFDVAVGIGLILTGAYLDIKALDKNAPKAAFKWTWCKGMMLAGAVT 120
|||||
QY 121 MAVTRLTEIIIPFTFANSYNRKLNLSNIAFGGFEPSFDINMQASALGFELSFKKSY 178
|||||
Db 121 MAVTRLTEIIIPFTFANSYNRKLNLSNIAFGGFEPSFDINMQASALGFELSFKKSY 178
|||||

RESULT 4

US-09-153-447-23

; Sequence 23, Application US/09153447

; GENERAL INFORMATION:

; APPLICANT: BERGSTROM, Sven

; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA

; FILE REFERENCE: 454312-3130.1

; CURRENT APPLICATION NUMBER: US/09/153,447

; CURRENT FILING DATE: 1998-09-15

; EARLIER APPLICATION NUMBER: 60/059,036

; EARLIER FILING DATE: 1997-09-10

; EARLIER APPLICATION NUMBER: 1041/97

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 23

; LENGTH: 177

; TYPE: PRT

; ORGANISM: Borrelia garinii

US-09-153-447-23

Query Match

Best Local Similarity 87.0%; Score 786.5; DB 15; Length 177;

Matches 154; Conservative 7; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNKFLIVLLAFVFSFAQADDSKSAFNLAGGKLLAYETSKKDPVLPVLLNLFGLFGI 60
|||||
Db 1 MNKFLIVLLAFVFSFAQADDSKSAFNLAGGKLLAYETSKKDPVLPVLLNLFGLFGI 59
|||||

QY 61 GSAQGDILGGFLILGFDVAVGIGLILTGAYLDIKALDKNAPKAAFKWTWCKGMMLAGAVT 120
|||||
Db 61 GSAQGDILGGFLILGFDVAVGIGLILTGAYLDIKALDKNAPKAAFKWTWCKGMMLAGAVT 119
|||||

QY 121 MAVTRLTEIIIPFTFANSYNRKLNLSNIAFGGFEPSFDINMQASALGFELSFKKSY 178
|||||

Db 120 MAVTRLTEIIVPFTFANNYNRKLNLSNIALGGFEPSFDINMQASALGFELSFKKSY 177
|||||

RESULT 5

US-09-156-447-23

; Sequence 23, Application US/09156447

; GENERAL INFORMATION:

; APPLICANT: BERGSTROM, Sven

; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA

; FILE REFERENCE: 454312-3130.1

; CURRENT APPLICATION NUMBER: US/09/156,447

; CURRENT FILING DATE: 1998-09-15

; PRIOR APPLICATION NUMBER: 60/059,036

; PRIOR FILING DATE: 1997-09-10

; PRIOR APPLICATION NUMBER: 1041/97

; PRIOR FILING DATE: 1997-09-10

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 23

; LENGTH: 177

; TYPE: PRT

; ORGANISM: Borrelia garinii

US-09-156-447-23

```
Query Match      87.0%; Score 786.5; DB 15; Length 177;
Best Local Similarity 86.5%; Pred. No. 2.6e-79;
Matches 154; Conservative 7; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNKFLIVVLLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKDPIVPELLNLFGLFGI 60
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MNKFLIFILVIFCAFSSFAQ-DDSKSTFNLGAGEKFLVYETNKKDSLVPELLNLFGLFGI 59
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GSFAQGDILGGFLILGFDVAGIGLILTGAYLIDIKALDKNAPKAAFKWTWCKGMMLAGAVT 120
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 60 GSFAQGDILGSLILGFDVAGIGLILTGAYLIDIKDFDNNAKKADFKWTWCKGMMLAGVVT 119
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 MAVTRLEIIPFTFANSYNRKLNLSNIAFGGPEPSFDINMGQASALGFELSFKKS 178
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 120 MAVTRLEIVLPFTFANNYNRKLNLSNIALGGPEPSFDINMGQASALGFELSFKKS 177
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
US-09-508-487-23
; Sequence 23, Application US/09508487
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3150
; CURRENT APPLICATION NUMBER: US/09/508,487
; CURRENT FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/059,036
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 1041/97
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: PCT/IB98/01424
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Borrelia garinii
US-09-508-487-23

Query Match      87.0%; Score 786.5; DB 19; Length 177;
Best Local Similarity 86.5%; Pred. No. 2.6e-79;
Matches 154; Conservative 7; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNKFLIVVLLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKDPIVPELLNLFGLFGI 60
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MNKFLIFILVIFCAFSSFAQ-DDSKSTFNLGAGEKFLVYETNKKDSLVPELLNLFGLFGI 59
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GSFAQGDILGGFLILGFDVAGIGLILTGAYLIDIKALDKNAPKAAFKWTWCKGMMLAGAVT 120
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 60 GSFAQGDILGSLILGFDVAGIGLILTGAYLIDIKDFDNNAKKADFKWTWCKGMMLAGVVT 119
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 MAVTRLEIIPFTFANSYNRKLNLSNIAFGGPEPSFDINMGQASALGFELSFKKS 178
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 120 MAVTRLEIVLPFTFANNYNRKLNLSNIALGGPEPSFDINMGQASALGFELSFKKS 177
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
US-09-153-447-19
; Sequence 19, Application US/09153447
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3130.1
; CURRENT APPLICATION NUMBER: US/09/153,447
; CURRENT FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: 60/059,036
; EARLIER FILING DATE: 1997-09-10
; EARLIER APPLICATION NUMBER: 1041/97
; EARLIER FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
```

```
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-153-447-19

Query Match      85.1%; Score 769.5; DB 15; Length 179;
Best Local Similarity 84.9%; Pred. No. 2.1e-77;
Matches 152; Conservative 10; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNKFLIVVLLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKDPIVPELLNLFGLFGI 59
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MNKLLIEVLATFCVFSFAQANDSKNGAFGMSAGEKLLVYETSKQDPIVPELLNLFGLFGI 60
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 60 IGSAQGDILGGFLILGFDVAGIGLILTGAYLIDIKALDKNAPKAAFKWTWCKGMMLAGAV 119
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 IGSAQGDILGSLILGFDVAGIGLILTGAYLIDIKALDITTKAFAFTWCKGMMLAGVV 120
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 120 TMVTRLEIIPFTFANSYNRKLNLSNIAFGGPEPSFDINMGQASALGFELSFKKS 178
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 TMVTRLEIIPFTFANSYNRKLNLSNIALGGPEPSFDVAMGQSSALGFELSFKKS 179
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
US-09-156-447-19
; Sequence 19, Application US/09156447
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3130.1
; CURRENT APPLICATION NUMBER: US/09/156,447
; CURRENT FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/059,036
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 1041/97
; PRIOR FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-156-447-19

Query Match      85.1%; Score 769.5; DB 15; Length 179;
Best Local Similarity 84.9%; Pred. No. 2.1e-77;
Matches 152; Conservative 10; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNKFLIVVLLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKDPIVPELLNLFGLFGI 59
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MNKLLIEVLATFCVFSFAQANDSKNGAFGMSAGEKLLVYETSKQDPIVPELLNLFGLFGI 60
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 60 IGSAQGDILGGFLILGFDVAGIGLILTGAYLIDIKALDKNAPKAAFKWTWCKGMMLAGAV 119
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 IGSAQGDILGSLILGFDVAGIGLILTGAYLIDIKALDITTKAFAFTWCKGMMLAGVV 120
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 120 TMVTRLEIIPFTFANSYNRKLNLSNIAFGGPEPSFDINMGQASALGFELSFKKS 178
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 TMVTRLEIIPFTFANSYNRKLNLSNIALGGPEPSFDVAMGQSSALGFELSFKKS 179
      ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
US-09-508-487-19
; Sequence 19, Application US/09508487
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3150
; CURRENT APPLICATION NUMBER: US/09/508,487
; CURRENT FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/059,036
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 1041/97
; PRIOR FILING DATE: 1997-09-10
```

; PRIOR APPLICATION NUMBER: PCT/IB98/01424
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-508-487-19

Query Match 85.1%; Score 769.5; DB 19; Length 179;
Best Local Similarity 84.9%; Pred. No. 2.1e-77;
Matches 152; Conservative 10; Mismatches 16; Indels 1; Gaps 1;
QY 1 MNKFLIVLLAFVCSFAQADDSKSAFNLGAGEKLLAYETSKKDPVIVPFLNLFGLFG 59
DB 1 MNKLLIFVLATFCVCSFAQADSKNGAFGMSAGEKLLVYETSKDPVIVPFLNLFGLFG 60
QY 60 IGSFAQGDILGGLIIGFDVAVGIGLITGAYLDIKALDNKPAKPKWTWKGMMLAGAV 119
DB 61 IGSFAQGDILGGLIIGFDVAVGIGLITGAYLDIKALDITTKAFAQWTWKGVMLAGVV 120
QY 120 TMAVTRLTETIIPTFANSYNNKLNATFGGPEPSPDINMGQASALGFELSPKKS 178
DB 121 TMAVTRLTETIIPTFANSYNNKLNATFGGPEPSPDINMGQASALGFELSPKKS 179

RESULT 10
US-09-153-447-31
; Sequence 31, Application US/09153447
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3130.1
; CURRENT APPLICATION NUMBER: US/09/153,447
; PRIOR FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: 60/059,036
; EARLIER FILING DATE: 1997-09-10
; EARLIER APPLICATION NUMBER: 1041/97
; EARLIER FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-153-447-31

Query Match 17.7%; Score 160; DB 15; Length 95;
Best Local Similarity 44.4%; Pred. No. 2.1e-09;
Matches 36; Conservative 12; Mismatches 31; Indels 2; Gaps 2;
QY 1 MNKFLIVLLAFVCSFAQADDSKSAFNLGAGEKLLAYETSKKDPVIVPFLNLFGLFG 60
DB 1 MKKIFTLILIFGLTIEIFA-TKDTQNRIEKGI-ESFNKYDKKKNPIGPFLLNLFPLFGI 58
QY 61 GSFAQGDILGGLIIGFDVAVG 81
DB 59 GSFVQGDYIGGGSVLGFNLIG 79

RESULT 11
US-09-156-447-31
; Sequence 31, Application US/09156447
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3130.1
; CURRENT APPLICATION NUMBER: US/09/156,447
; CURRENT FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/059,036
; PRIOR FILING DATE: 1997-09-10
; CURRENT APPLICATION NUMBER: 1041/97

; PRIOR FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-156-447-31

Query Match 17.7%; Score 160; DB 15; Length 95;
Best Local Similarity 44.4%; Pred. No. 2.1e-09;
Matches 36; Conservative 12; Mismatches 31; Indels 2; Gaps 2;
QY 1 MNKFLIVLLAFVCSFAQADDSKSAFNLGAGEKLLAYETSKKDPVIVPFLNLFGLFG 60
DB 1 MKKIFTLILIFGLTIEIFA-TKDTQNRIEKGI-ESFNKYDKKKNPIGPFLLNLFPLFGI 58
QY 61 GSFAQGDILGGLIIGFDVAVG 81
DB 59 GSFVQGDYIGGGSVLGFNLIG 79

RESULT 12
US-09-508-487-31
; Sequence 31, Application US/09508487
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3150
; CURRENT APPLICATION NUMBER: US/09/508,487
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/059,036
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 1041/97
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: PCT/IB98/01424
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-508-487-31

Query Match 17.7%; Score 160; DB 19; Length 95;
Best Local Similarity 44.4%; Pred. No. 2.1e-09;
Matches 36; Conservative 12; Mismatches 31; Indels 2; Gaps 2;
QY 1 MNKFLIVLLAFVCSFAQADDSKSAFNLGAGEKLLAYETSKKDPVIVPFLNLFGLFG 60
DB 1 MKKIFTLILIFGLTIEIFA-TKDTQNRIEKGI-ESFNKYDKKKNPIGPFLLNLFPLFGI 58
QY 61 GSFAQGDILGGLIIGFDVAVG 81
DB 59 GSFVQGDYIGGGSVLGFNLIG 79

RESULT 13
US-09-791-537-91922
; Sequence 91922, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 91922

; LENGTH: 95
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-791-537-91922

Query Match 17.7%; Score 160; DB 21; Length 95;
Best Local Similarity 44.4%; Pred. No. 2.1e-09;
Matches 36; Conservative 12; Mismatches 31; Indels 2; Gaps 2;

Oy 1 MNKFLIVLLAFVFPSSFAQADDSKSAFNLGAGEKLLAYETSKKDPVFPFLNLFGLPFGI 60
Db 1 MKKIPTLILFGLTIEIFA-TKDTQNRKEGI-ESFNKYDKKKNPIGPFLLNLFPLPFGI 58

Oy 61 GSFAQGDILGGFLILGFDVAG 81
Db 59 GSFVQGDYIGGSGVLGFNLG 79

RESULT 14
US-09-153-447-1
; Sequence 1, Application US/09153447
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3130.1
; CURRENT APPLICATION NUMBER: US/09/153.447
; CURRENT FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: 60/059,036
; EARLIER FILING DATE: 1997-09-10
; EARLIER APPLICATION NUMBER: 1041/97
; EARLIER FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-153-447-1

Query Match 13.7%; Score 124; DB 15; Length 25;
Best Local Similarity 96.0%; Pred. No. 4.3e-06;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 41 TSKKDPVFPFLNLFGLFGIGSFAQ 65
Db 1 TSKQDPVFPFLNLFGLFGIGSFAQ 25

RESULT 15
US-09-156-447-1
; Sequence 1, Application US/09156447
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3130.1
; CURRENT APPLICATION NUMBER: US/09/156.447
; CURRENT FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/059,036
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 1041/97
; PRIOR FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-156-447-1

Query Match 13.7%; Score 124; DB 15; Length 25;
Best Local Similarity 96.0%; Pred. No. 4.3e-06;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 41 TSKKDPVFPFLNLFGLFGIGSFAQ 65
Db 1 TSKQDPVFPFLNLFGLFGIGSFAQ 25

Search completed: March 14, 2003, 14:35:07
Job time : 146.333 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:25:51 ; Search time 10 Seconds
(without alignments)
820.439 Million cell updates/sec

Title: US-09-508-487-21

Perfect score: 904

Sequence: 1 MNRFLIVLLAFCVFSSFAQ.....DINMGQASALGFELSPKSY 178

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|------------------------|
| 1 | 97 | 10.7 | 417 | 9 | US-09-738-626-6850 |
| 2 | 83.5 | 9.2 | 650 | 10 | US-09-815-242-13341, A |
| 3 | 81.5 | 9.0 | 413 | 9 | US-09-910-186A-22 |
| 4 | 81.5 | 9.0 | 852 | 9 | US-10-011-588-25 |
| 5 | 78.5 | 8.7 | 663 | 10 | US-09-815-242-10070 |
| 6 | 78.5 | 8.7 | 663 | 10 | US-09-815-242-14080 |
| 7 | 76 | 8.4 | 411 | 10 | US-09-815-242-5875 |
| 8 | 76 | 8.4 | 432 | 10 | US-09-815-242-13058 |
| 9 | 75.5 | 8.4 | 206 | 10 | US-09-815-242-11462 |
| 10 | 73.5 | 8.1 | 1161 | 8 | US-08-910-386A-20 |
| 11 | 72.5 | 8.0 | 451 | 10 | US-09-741-669-422 |
| 12 | 72 | 8.0 | 452 | 10 | US-09-925-637-16 |
| 13 | 71.5 | 7.9 | 220 | 10 | US-09-881-752A-16 |
| 14 | 71.5 | 7.9 | 292 | 9 | US-09-738-626-4426 |
| 15 | 71.5 | 7.9 | 354 | 9 | US-09-860-670-134 |
| 16 | 71.5 | 7.9 | 922 | 10 | US-09-886-468-19 |
| 17 | 71 | 7.9 | 406 | 10 | US-09-215-450-26 |
| 18 | 70.5 | 7.8 | 456 | 10 | US-09-815-242-4932 |
| 19 | 70.5 | 7.8 | 463 | 10 | US-09-815-242-10662 |

| | | | | | |
|----|------|-----|------|----|---------------------|
| 20 | 70 | 7.7 | 332 | 9 | US-09-738-626-5123 |
| 21 | 69.5 | 7.7 | 337 | 10 | US-09-815-242-11246 |
| 22 | 69.5 | 7.7 | 487 | 9 | US-09-738-626-4875 |
| 23 | 69.5 | 7.7 | 513 | 9 | US-09-738-626-6860 |
| 24 | 69 | 7.6 | 346 | 9 | US-09-738-626-5811 |
| 25 | 69 | 7.6 | 446 | 9 | US-10-217-096-4 |
| 26 | 69 | 7.6 | 451 | 10 | US-09-815-242-5633 |
| 27 | 69 | 7.6 | 451 | 10 | US-09-815-242-12384 |
| 28 | 69 | 7.6 | 2285 | 10 | US-09-932-183A-2 |
| 29 | 68.5 | 7.6 | 431 | 9 | US-09-738-626-4697 |
| 30 | 68 | 7.5 | 407 | 10 | US-09-815-242-11180 |
| 31 | 68 | 7.5 | 450 | 10 | US-09-815-242-11384 |
| 32 | 67 | 7.4 | 296 | 10 | US-09-815-242-11212 |
| 33 | 67 | 7.4 | 383 | 9 | US-09-895-913A-152 |
| 34 | 67 | 7.4 | 451 | 9 | US-09-925-179-68 |
| 35 | 67 | 7.4 | 469 | 10 | US-09-815-242-5281 |
| 36 | 67 | 7.4 | 497 | 10 | US-09-815-242-12626 |
| 37 | 67 | 7.4 | 643 | 10 | US-09-815-242-5394 |
| 38 | 67 | 7.4 | 657 | 10 | US-09-815-242-12301 |
| 39 | 66.5 | 7.4 | 844 | 10 | US-09-272-809-6 |
| 40 | 66.5 | 7.4 | 933 | 10 | US-09-815-242-11817 |
| 41 | 66 | 7.3 | 380 | 10 | US-09-815-242-10609 |
| 42 | 66 | 7.3 | 544 | 9 | US-09-738-626-3927 |
| 43 | 66 | 7.3 | 544 | 10 | US-09-948-777-2 |
| 44 | 66 | 7.3 | 702 | 10 | US-09-815-242-11858 |
| 45 | 66 | 7.3 | 2042 | 9 | US-10-192-584-6 |

ALIGNMENTS

RESULT 1

US-09-738-626-6850
; Sequence 6850, Application US/09738626
; Publication No. US20020197605A1

GENERAL INFORMATION:

; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6850
; LENGTH: 417
; TYPE: PPT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6850

Query Match 10.7%; Score 97; DB 9; Length 417;
Best Local Similarity 23.1%; Pred. No. 0.022;
Matches 50; Conservative 24; Mismatches 62; Indels 80; Gaps 9;

QY 5 LIVVLLAFVSSFAQADDSKAFNLGAGEKLLAVETSKK-----DPI 47
DB 135 MTAALLAF-----LGLGMAVIKSDTLFKVTRERLVVVKTTITAFVPL 178
QY 48 VP-FLNLFGLFGISFAQGDIL-----GGFLILGFDVAGIGLITGAYLDIKALDKNA 100

NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant protein encoded by SEQ ID NO:24
US-10-011-588-25

Query Match 9.0%; Score 81.5; DB 9; Length 852;
Best Local Similarity 23.8%; Pred. No. 2.2;
Matches 36; Conservative 24; Mismatches 56; Indels 35; Gaps 6;
Qy 47 IVPPELLNLFGLGSGFAQGDILGGLILGFDV-----GIGLILGAYLD----- 92
Db 619 IVPY-IGIALNVG-NETAKGFENAFETAGASILLEFIPELLIPVVGAFLESYIDNKNK 676
Qy 93 -IKALDKNAPKAFAKWTGKGMLAGAVTMAVTRL-----TEIIIPFT 134
Db 677 IIKTIDNALTNRNEKWSMDYGLIWAQLSTVNTQFYTIKEGMYKALNYQAQALEEIIKYR 736
Qy 135 FANSYNRKLKNSLNIAFGFSPEDINMGQA 165
Db 737 Y-NIYSEKSNINIDFNDINSKLNIGINQA 766

RESULT 5
US-09-815-242-10070
; Sequence 10070, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10070
; LENGTH: 663
; TYPE: PRT
; ORGANISM: Salmonella typhi
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(663)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-10070

Query Match 8.7%; Score 78.5; DB 10; Length 663;
Best Local Similarity 22.6%; Pred. No. 3.3;
Matches 38; Conservative 19; Mismatches 72; Indels 39; Gaps 6;
Qy 28 FNLGAGEKLLAYETSKDPPIVFFLLNLF-----LGFGISFAQGDILG----- 70
Db 118 FVIGLMLNVPLQIGARDVAFPPFLNLSFWFTVGVILVNSLGVGEFAQTGLAYPPPLS 177
Qy 71 -----GFLILGFDVAVGIGLITGAYLDIKALDKNAP-KAAFK---WTW-----GKG 112
Db 178 GIEYSPGVGVYDIWISLQSLGIGTTLTGINFVITILKMRAPGTMFKMPVFTWASLCANV 237
Qy 113 MMLAG-----AVTMAVTRLTEIIIPFTFANSYNRKLKNSLNIAFGGFEP 156
Db 238 LIIASFPILTVTVALLTDRLYLGHFFFTNDMGNNMMYINLIWAGHP 285

RESULT 6
US-09-815-242-14080
; Sequence 14080, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14080
; LENGTH: 663
; TYPE: PRT
; ORGANISM: Salmonella typhi
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(663)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-14080

Query Match 8.7%; Score 78.5; DB 10; Length 663;
Best Local Similarity 22.6%; Pred. No. 3.3;
Matches 38; Conservative 19; Mismatches 72; Indels 39; Gaps 6;
Qy 28 FNLGAGEKLLAYETSKDPPIVFFLLNLF-----LGFGISFAQGDILG----- 70
Db 118 FVIGLMLNVPLQIGARDVAFPPFLNLSFWFTVGVILVNSLGVGEFAQTGLAYPPPLS 177
Qy 71 -----GFLILGFDVAVGIGLITGAYLDIKALDKNAP-KAAFK---WTW-----GKG 112
Db 178 GIEYSPGVGVYDIWISLQSLGIGTTLTGINFVITILKMRAPGTMFKMPVFTWASLCANV 237
Qy 113 MMLAG-----AVTMAVTRLTEIIIPFTFANSYNRKLKNSLNIAFGGFEP 156
Db 238 LIIASFPILTVTVALLTDRLYLGHFFFTNDMGNNMMYINLIWAGHP 285

RESULT 7
US-09-815-242-10070
; Sequence 10070, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10070
; LENGTH: 663
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10070

Query Match 8.7%; Score 78.5; DB 10; Length 663;
Best Local Similarity 22.6%; Pred. No. 3.3;
Matches 38; Conservative 19; Mismatches 72; Indels 39; Gaps 6;
Qy 28 FNLGAGEKLLAYETSKDPPIVFFLLNLF-----GIGSFAQGDILG----- 70


```
Best Local Similarity 23.6%; Pred. No. 1.6;
Matches 33; Conservative 20; Mismatches 38; Indels 49; Gaps 6;

Qy 61 GSFAQGDILGFLGFDVAGIGILITGAYLDIKALDNKAPKAFKWTG-----K 111
Db 86 GHFVOGHI-----DAIGV-----IEKIIHNANQVDFISASETEILLLCVEQ 126
Qy 112 GMLAGAVTMAVTRLTE-----IIIPTEFANSYNR--KLKNSLNTAFG----- 152
Db 127 GSIADVGSULTSKVEKGFWLTIPTLTLEFKAYLKRVRNIETDMLVRSVASILKK 186
Qy 153 --GFEPSPDINMGQASALGF 170
Db 187 TKGEKFNFSWNEADALTGY 206

RESULT 10
US-08-910-386A-20
; Sequence 20, Application US/08910386A
; Patent No. US20020092041A1
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Hulbert, Scot
; APPLICANT: Richter, Todd
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,386A
; FILING DATE: 13-AUG-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-05895005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1161 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-910-386A-20

Query Match 8.1%; Score 73.5; DB 8; Length 1161;
Best Local Similarity 28.4%; Pred. No. 22;
Matches 31; Conservative 17; Mismatches 42; Indels 19; Gaps 5;

Qy 76 GFDAVGIGILITGAYLDIKALDNKAPKAFKWT--WCKGMMLAGAVTMAVTRLTEIIPF 133
Db 830 GFNSIETASITSAAIYVLL---ALIVLFYTRKNWPNRSRVAGSTRKEVTFTEVPVL 886
Qy 134 TFAN-----SYNRKLKNSLNTAFGFGPEPSFDINMGQASALGFELSPKK 176
Db 887 TFENVVRATGSFNA-----SNCIGSGGFGATYKAEI----APGFLVAVKR 927
```

```
RESULT 11
US-09-741-669-422
; Sequence 422, Application US/09741669
; Patent No. US20020022718A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlson, Kari L.
; APPLICANT: Zyskind, Judith W.
; TITLE OF INVENTION: Genes identified as required for
; TITLE OF INVENTION: proliferation of E. coli
; FILE REFERENCE: ELITRA.009A
; CURRENT APPLICATION NUMBER: US/09/741.669
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 422
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-09-741-669-422

Query Match 8.0%; Score 72.5; DB 10; Length 451;
Best Local Similarity 26.2%; Pred. No. 8.7;
Matches 34; Conservative 14; Mismatches 45; Indels 37; Gaps 7;

Qy 46 PIVFPLMLFLGFGIGSFAQGDILGGLILGFDVAGIGILITGAYLD-----IKAL----- 96
Db 19 PIVIIIFSKILG-----MKAGDCFKAGLHIGIFGVGILVI-GLMLDSIGSPAAMAEKF 72
Qy 97 DKNAPKAAFKW-----TWGKGMLAGA-----VTMAVTRLTEIIPFTEFANSYNRKL 143
Db 73 DLNLHVVDVGPWGSSPMTWASQIALVPIALVNVAMLLTRMTRV-----NVDI 123
Qy 144 KNSLNIAPFG 153
Db 124 WNIWMTFTG 133

RESULT 12
US-09-925-637-16
; Sequence 16, Application US/09925637
; Patent No. US2002010338A1
; GENERAL INFORMATION:
; APPLICANT: Choi
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypepti
; FILE REFERENCE: PB560
; CURRENT APPLICATION NUMBER: US/09/925.637
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/151,933
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 08/781,986
; PRIOR FILING DATE: 1997-01-03
; PRIOR APPLICATION NUMBER: US 08/956,171
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/009,861
; PRIOR FILING DATE: 1996-01-06
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-925-637-16

Query Match 8.0%; Score 72; DB 10; Length 452;
Best Local Similarity 21.3%; Pred. No. 9.9;
Matches 37; Conservative 28; Mismatches 71; Indels 38; Gaps 6;

Qy 27 AFNLGA-GEKLLAYETSKKDFIVPFLNLFLGFGI---GSFAQGDILGFLGFDVAGI 82
```

Db 24 AFLGIRYGGVLAHNGKHP-----RVLVGRDTRVSGEMLESALIGLISICAEVMRL 77
QY 83 GLILT-----GAYLDIKALDKNAPKA--AFKWTWCKMMLAGAVTMVTRLTEII 130
Db 78 GIISTPGVAYLTRDMGAELGVMISASHNPVADNGIKFFGSDGPKLSDEQNEIEALLDQE 137
QY 131 IP-----FTFANSYNRKLNLSNIAFGGFEPSFDINMGQASAL 168
Db 138 NPELPRPVGNDIVHYSIDYEGAQKYLKSTVDVNFEGLIKIALDANGSTSSL 191

RESULT 13
US-09-881-752A-16
; Sequence 16, Application US/09881752A
; Patent No. US20020115078A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/041002
; CURRENT APPLICATION NUMBER: US/09/881,752A
; PRIOR FILING DATE: 2001-06-15
; PRIOR FILING DATE: 1997-04-01
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-881-752A-16

Query Match 7.9%; Score 71.5; DB 10; Length 220;
Best Local Similarity 26.1%; Pred. No. 4.6;
Matches 47; Conservative 20; Mismatches 62; Indels 51; Gaps 10;
QY 1 MNKFLIVLLAFVCFVSFAQD---DSKS-----AFNLGAGEKLLAYETSKDPIV 48
Db 1 MNKTTVKILMGMAELSLQAAEAELDEKSKPKFADRNFTYLGVGQYQLSAINTSFSTESV 60
QY 49 PFLNLFL---GFGIGSFAQGDILGFLILGFDV---GIGLILT-----GAYLDIK 94
Db 61 D--KSYFMTGNGFV-----VLGGKFVAKTQAVEHVGFRYGLFYDQTFSSHKSIYI 111
QY 95 ALD-----KNAPKAAFKWTWCKMMLAGAVTMVTRLTEIIPFTFANSYNRKLNLSL 147
Db 112 GLEFGSLWDAFNPKPMFLGLEFLG---IAGATYMPGGAMHGII-----AQNLRKNSL 162

RESULT 14
US-09-738-626-4426
; Sequence 4426, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OKAZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626

; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent in ver. 3.0
; SEQ ID NO 4426
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4426

Query Match 7.9%; Score 71.5; DB 9; Length 292;
Best Local Similarity 26.5%; Pred. No. 6.5;
Matches 43; Conservative 18; Mismatches 56; Indels 45; Gaps 9;
QY 18 FAQADDSKSAF-----NLGAGEKLLAYETSKDPIVPELNLFLGFGISFAQGDIL-- 69
Db 18 FVSASDPQSIIRSEPRADRCYGRKLLA-QLNPTWPTP-----ICQFALNRSVPA 66
QY 70 --GGFLLGFDVAGIGLILTGAYLDIKALDKNAPKAAFKWTWCKMMLAGAVTMVTRLT 127
Db 67 SANEFYIAGFP--GITIIQT-VLEDVTSLSKLNPL-----LRSVPATDVIYFAVN 114
QY 128 EIIIPTEANSYNRKLNLSNIAF-----GGFEPSF 158
Db 115 EETTLGGFAHIYNGEIKRSF-IAYERVPEDNGIPGGFETPY 155

RESULT 15
US-09-860-670-134
; Sequence 134, Application US/09860670
; Patent No. US20020165137A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA137P1
; CURRENT APPLICATION NUMBER: US/09/860,670
; CURRENT FILING DATE: 2001-05-21
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 134
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-860-670-134

Query Match 7.9%; Score 71.5; DB 9; Length 354;
Best Local Similarity 22.1%; Pred. No. 8.3;
Matches 36; Conservative 19; Mismatches 49; Indels 59; Gaps 5;
QY 1 MNKFLIVLLAFVCFVSFAQADDSKSAFNIGAGEKLLAYETSKDPIVPELNL-FLGFG 59
Db 196 MGAFLAGVLLASSEYRHALES-----IEPFKGLLLGLFFIGVG 234
QY 60 I-----GSFAQGDILGFLILGFDVAGIGLILTGAYLDIKALDKNAPKAAFKW----- 107
Db 235 MSIDFGTLLLENPLRIVILLIGF-----LIKTAMLWLITARLPQVKNQRWFAVLIGQG 288
QY 108 -----TWGKMMLAGAVTMVTRLTEI 131
Db 289 SEFAFVVFGAAQMANVLEPEWAKSLTVALVSMATPILLVIL 331

Search completed: March 14, 2003, 14:37:05
Job time : 12 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: March 14, 2003, 14:22:16 ; Search time 12.264 Seconds
(without alignments)
1387.454 Million cell updates/sec

Title: US-09-508-487-23
Perfect score: 908
Sequence: 1 MNKFLFILVIFCAFSSFAQ.....DINMGQASALGFLGFSFKSY 177

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 737 | 81.2 | 179 | 2 B70104 | hypothetical prote |
| 2 | 309 | 34.0 | 161 | 2 A70207 | conserved hypothet |
| 3 | 301.5 | 33.2 | 170 | 2 F70243 | conserved hypothet |
| 4 | 240 | 26.4 | 190 | 2 G70239 | conserved hypothet |
| 5 | 93 | 10.2 | 779 | 2 AD1457 | ATP dependent heli |
| 6 | 88.5 | 9.7 | 376 | 2 E95879 | probable sugar ABC |
| 7 | 88.5 | 9.7 | 427 | 2 S55905 | phosphotransferase |
| 8 | 88.5 | 9.7 | 451 | 2 C64376 | gacC protein - Esc |
| 9 | 88.5 | 9.7 | 451 | 2 G90990 | PTS system galacti |
| 10 | 88.5 | 9.7 | 451 | 2 A85836 | PTS system galacti |
| 11 | 87 | 9.6 | 1291 | 1 A48940 | bontoxilysin (EC 3 |
| 12 | 86.5 | 9.5 | 389 | 2 T28954 | hypothetical prote |
| 13 | 86 | 9.5 | 272 | 2 H75011 | hypothetical prote |
| 14 | 86 | 9.5 | 406 | 1 JC5041 | fosmidomycin resis |
| 15 | 86 | 9.5 | 406 | 2 D90695 | fosmidomycin resis |
| 16 | 86 | 9.5 | 406 | 2 H85545 | fosmidomycin resis |
| 17 | 85 | 9.4 | 235 | 2 D69597 | probable holoctoc |
| 18 | 85 | 9.4 | 692 | 2 H70362 | glycogen phosphory |
| 19 | 84.5 | 9.3 | 457 | 2 AH0898 | PTS system, galact |
| 20 | 84 | 9.3 | 406 | 2 AH0563 | fosmidomycin resis |
| 21 | 83.5 | 9.2 | 449 | 2 S37867 | hypothetical prote |
| 22 | 82 | 9.0 | 357 | 2 H70346 | undecaprenyl-phosp |
| 23 | 81.5 | 9.0 | 206 | 2 F64716 | riboflavin synthas |
| 24 | 81.5 | 9.0 | 394 | 2 T13721 | NADH2 dehydrogenas |
| 25 | 81.5 | 9.0 | 1185 | 2 AG2457 | transcription-repa |
| 26 | 81 | 8.9 | 174 | 2 T11491 | NADH2 dehydrogenas |
| 27 | 81 | 8.9 | 371 | 2 S55588 | D-nopaline dehydro |
| 28 | 80.5 | 8.9 | 183 | 2 H72250 | conserved hypothet |
| 29 | 80.5 | 8.9 | 315 | 2 F70203 | xylose operon regu |

30 80 8.8 355 2 A97271 glycosyltransferas
31 80 8.8 626 2 A70169 hypothetical prote
32 80 8.8 1291 2 I40631 non-proteolytic bo
33 79.5 8.8 436 2 D83498 C4-dicarboxylate t
34 79.5 8.8 893 2 A37284 surface-array prot
35 79 8.7 414 2 B97739 hypothetical prote
36 79 8.7 469 2 H97064 probable sugar-pro
37 79 8.7 560 2 T32727 hypothetical prote
38 79 8.7 574 2 D70114 PTS system, maltos
39 78.5 8.6 430 2 H69864 macrolide-efflux p
40 78.5 8.6 473 2 H97092 closely related to
41 78.5 8.6 936 1 I40705 bacterial adhesin
42 78.5 8.6 939 2 I41197 eae protein (enter
43 78 8.6 264 2 S73384 hypothetical prote
44 78 8.6 379 2 E88109 protein T24E12.8 |
45 78 8.6 488 2 H90423 sugar transport pr

ALIGNMENTS

RESULT 1

B70104

hypothetical protein BB0034 - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Mar-2000

C:Accession: B70104

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; W

son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; V

; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943; PMID:9403685

A:Accession: B70104

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-179 <KLE>

A:Cross-references: GB:AE001117; GB:AE000783; NID:92687907; PIDN:AAC66426.1; PID:926

A:Experimental source: strain B31

C:Superfamily: Lyme disease spirochete plasmid hypothetical protein BBA01

Query Match 81.2%; Score 737; DB 2; Length 179;

Best Local Similarity 81.0%; Pred. No. 5.5e-60;

Matches 145; Conservative 14; Mismatches 18; Indels 2; Gaps 2;

Qy 1 MNKFLFILVIFCAFSSFAQ-DDSKS-TFNLGAGEKFLVYETNKKDSLVPFLNLFGLFG 58

Db 1 MNKLLIFVLATFCVFSPPAQANDSKNGAFGMSAGEKLLVYETSKQDPIVPFLNLFGLFG 60

Qy 59 IGSAQGDILGSLILGFDVAVGIGLITGAYLDIKDFDNNAKKADFKWTGKGWMLAGVV 118

Db 61 IGSAQGDILGSLILGFDVAVGIGLITGAYLDIKALDITKKAAFOWTGKGWMLAGVV 120

Qy 119 TMAYTRTEIILPTFANNYRNKLNLSNIALGGPEPSFDINMGQASALGFLGFSFKSY 177

Db 121 TMAYTRTEIILPTFANNYRNKLNLSNIALGGPEPSFDVAMGQSSALGFLGFSFKSY 179

RESULT 2

A70207

conserved hypothetical protein BBA01 - Lyme disease spirochete plasmid A/lp54

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Mar-2000

C:Accession: A70207

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; W

son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; V

; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943; PMID:9403685

A:Accession: A70207

A;Residues: 36-217, 'G', 219-224, 'S', 226-246 <SZ>
A;Cross-references: EMBL:Z11934; NID:g40384; PIDN:CAA77991.1; PID:g40384
R;Kurazono, H.; Mochida, S.; Binz, T.; Eisell, U.; Quanz, M.; Grebenstein, O.; Wernars
J. Biol. Chem. 267, 14721-14729, 1992
A;Title: Minimal essential domains specifying toxicity of the light chains of tetanus
A;Reference number: A42871; MUID:92340509; PMID:1634516
A;Accession: A42871
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-313, 'S', 315-451 <KUR>
A;Experimental source: strain Okra
A;Note: sequence extracted from NCBI backbone (NCBIP:109365)
R;DasGupta, B.R.; Datta, A.
Biochimie 70, 811-817, 1988
A;Title: Botulinum neurotoxin type B (strain 657): partial sequence and similarity w
A;Reference number: S07155; MUID:89000987; PMID:3139097
A;Accession: S07155
A;Molecule type: protein
A;Residues: 2-29, 'M', 31-45 <DAS>
A;Accession: S08562
A;Molecule type: protein
A;Residues: 442-463, 'R', 465-467 <DA2>
R;Schmidt, J.J.; Sathyamoorthy, V.; DasGupta, B.R.
Arch. Biochem. Biophys. 238, 544-548, 1985
A;Title: Partial amino acid sequences of botulinum neurotoxins types B and E.
A;Reference number: S07128; MUID:85197963; PMID:3888113
A;Accession: S07128
A;Status: preliminary
A;Molecule type: protein
A;Residues: 2-16 <SCH1>
A;Accession: S08573
A;Status: preliminary
A;Molecule type: protein
A;Residues: 2-17 <SCH2>
A;Accession: S08574
A;Status: preliminary
A;Molecule type: protein
A;Residues: 442-459 <SCH3>
R;Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta,
Nature 359, 832-835, 1992
A;Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by prote
A;Reference number: S21275; MUID:93063293; PMID:1331807
A;Contents: annotation
C;Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic sy
C;Genetics:
A;Gene: bont/b
C;Function:
A;Description: catalyzes hydrolysis of a Gln-Phe peptide bond in synaptobrevin 2
C;Superfamily: tetanus toxin
C;Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
F;2-441/Product: bontoxilysin B light chain #status experimental <LIGHT>
F;442-1291/Product: bontoxilysin B heavy chain #status experimental <HVV>
F;230,234/Binding site: zinc (His) #status predicted
F;231/Active site: Glu #status predicted

| | Query Match | 9.68; | Score 87; | DB 1; | Length 1291; |
|----|---------------------------|--|----------------|----------------|--------------------|
| | Best Local Similarity | 22.88; | Pred. No. 7.5; | | |
| | Matches 38; | Conservative | 31; | Mismatches 58; | Indels 40; Gaps 7; |
| Qy | 37 | VYETNKDS-----LVPEFL---LNLFGLFGIGFAQG-DILGGSLILGPD---- | AV 79 | | |
| | : : : | | | | |
| Db | 602 | VIEANKSNWDMKIADISLIYPVIGIALNVGNETAQNFAFEAGASILLEFIPELLIP | 661 | | |
| | : : : | | | | |
| Qy | 80 | GIGIILTAYLD-----IKDPDNNAKKADFKWTKWGKMMLAGVVVMATRL----- | 125 | | |
| | : : : | | | | |
| Db | 662 | VVGAFLESYIDNNKIIKTIDNALTRNEKWSMDMYGLIIVAQWLSTVTQFYTIKEGM | 721 | | |
| | : : : | | | | |
| Qy | 126 | -----TEVLPTFTANNRKLNLSNALALGGFFPSFDINMQA | 164 | | |
| | : : : | | | | |
| Db | 722 | ALNQOAALEELIKRY-NIYSEKESNINDFNDINSKLNEGINQA | 767 | | |
| | : : : | | | | |

RESULT 12

RESULT 12

T28954

hypothetical protein F28A12.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000

C:Accession: T28954

R:Sammons, L.; Murray, J.

submitted to the EMBL Data Library, July 1996

A:Description: The sequence of C. elegans cosmid F28A12.

A:Reference number: #20547

A:Accession: T28954

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-389 <SAM>

A:Cross-references: EMBL:U64851; PIDN:AAC47989.1; GSPDB:GN00023; CESP:F28A12.4

A:Experimental source: strain Bristol N2; clone F28A12

C:Genetics:

A:Gene: CESP:F28A12.4

A:Map position: 5

A:Introns: 104/2; 147/3; 175/3; 258/3; 301/3; 341/3

C:Superfamily: pepsin

Query Match 9.5%; Score 86.5; DB 2; Length 389;

Best Local Similarity 24.7%; Pred. No. 2.3;

Matches 40; Conservative 22; Mismatches 57; Indels 43; Gaps 8;

QY 22 DSKSTFNLGAGEKFLVYETNKKDSLVPFLNLFGLFGIGSFAOGDILGGSLILGFDVAVGI 81

DB 108 EKKRLFNKASSTYIA--TNR-----PWQI-----KYASGDAYG---TLGIDTWKI 148

QY 82 G-----LILTGAYLDIKDFDNNAKADFKWTWKGKMLAGVVTWAVTRLTVEILPFTFA 135

DB 149 GGSGEAQLAIPRSYLGAVD---TVGSDFKWSPKEGIFGLAFTALAVDNITPPII----- 199

QY 136 NNNKLNLSNIALGGPEPSFDINMQASALGFGLSFKSKY 177

DB 200 -----NAINQGLLD-QPLFTTWFGQAGPATSASGFTY 232

RESULT 13

H75011

hypothetical protein PAB1278 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 28-Jul-2000

C:Accession: H75011

R:anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: Insights into archaeal chromosome stru

A:Reference number: A75001

A:Accession: H75011

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-272 <KAW>

A:Cross-references: GB:AJ248288; GB:AL096836; NID:g5458960; PIDN:CAB50534.1; PID:g545904

A:Experimental source: strain Orsay

C:Genetics:

C:Superfamily: Pyrococcus abyssi hypothetical protein PAB1278

Query Match 9.5%; Score 86; DB 2; Length 272;

Best Local Similarity 19.4%; Pred. No. 1.8;

Matches 37; Conservative 32; Mismatches 66; Indels 56; Gaps 6;

QY 3 KFLIFILVFCFAFFSFAODDSKSTENLGAGEKFLVYETNKK----- 43

DB 6 KALVIVIAIYATLRFASIDTLMSL-LGLGLKLGIRMSPNGFLIFGLIALGIMSDY 64

QY 44 ----DSLVPFLNLFGLFGIGSFAOGDILGGSLILGFDVAVGI-----GLILTGAYLDIK 93

DB 65 SVIVTSSIPFFLYSF----VWFQSNVILGSNVIASMGIEVSSSTLLGLIVLLAFSFTV 120

QY 94 DFDNNAKADFKW-TWKGKMLAGVVTWA-----VTRLTEIVLP 131

DB 121 DSYNSYERKLEVTWESYRFGLLIALSVILTSIMFLALSRIDLSTQPVYVKVILLVLLVVP 180

QY 132 FTFANNYNRKL 142

DB 181 FLLVQEGKKV 191

RESULT 14

JC5041

fosmidomycin resistance protein - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Mar-2002

C:Accession: JC5041; F64778

R:Fujisaki, S.; Ohnuma, S.; Horiuchi, T.; Takahashi, I.; Tsukui, S.; Nishimura, Y.;

Gene 175, 83-87, 1996

A:Title: Cloning of a gene from Escherichia coli that confers resistance to fosmidom

A:Reference number: JC5041; MUID:97074653; PMID:8917080

A:Accession: JC5041

A:Molecule type: DNA

A:Residues: 1-406 <FUJ>

A:Cross-references: DBJ:D73370; NID:g1019359; PIDN:BAA11120.1; PID:g1019360

A:Experimental source: strain DH5alpha

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

.A.; Rose, D.J.; Mau, B.; Shaio, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: F64778

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-406 <BLAT>

A:Cross-references: GB:AB000154; GB:U00096; NID:g1786683; PIDN:AAC73581.1; PID:g1786

A:Experimental source: strain K-12, substrain MGI655

C:Genetics:

A:Gene: fsr

C:Function:

A:Description: confers resistance to fosmidomycin

C:Superfamily: fosmidomycin resistance protein

C:Keywords: antibiotic resistance; transmembrane protein

F:104-120/Domain: transmembrane #status predicted <TM1>

F:181-197/Domain: transmembrane #status predicted <TM2>

F:225-241/Domain: transmembrane #status predicted <TM3>

F:267-283/Domain: transmembrane #status predicted <TM4>

F:294-310/Domain: transmembrane #status predicted <TM5>

F:320-336/Domain: transmembrane #status predicted <TM6>

F:358-374/Domain: transmembrane #status predicted <TM7>

F:381-397/Domain: transmembrane #status predicted <TM8>

Query Match 9.5%; Score 86; DB 1; Length 406;

Best Local Similarity 23.2%; Pred. No. 2.7;

Matches 39; Conservative 27; Mismatches 64; Indels 38; Gaps 6;

QY 6 FPLIVFCFAFFSFAODDSKSTENLGAGEKFLVYETNKKDSLVPFLNLFGLFGIGSFAOG 65

DB 231 ILLILIFSKYFYMASSISYTYFL--MQKFGLSIQNALHLFAFLF-----AVAAG 279

QY 66 DILGSLILGFDVAVGILITGAYLDIKDFDNNAKADFKWTWKGKMLAGVVTWAVTRL 125

DB 280 TWIGFV---GDKIGRKVYINGSLGVAPFTLILPYASLHWI-----GVLT----- 322

QY 126 TEIVLPFTFANNYNRKLNSLNIALGGPEPSFDINMQASALGFGLSF 173

DB 323 --VIIGFTLASAFSAILVYAQLLPG-----RIGMVSGLFFGFAF 360

RESULT 15

D90695

fosmidomycin resistance protein [imported] - Escherichia coli (strain O157:H7, subst

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C:Accession: D90695

R:Hayashi, T.; Mahi, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han,

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: D90695
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-406 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA033955.1; PID:g13359989; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECs0532
C:Superfamily: fosmidmycin resistance protein

Query Match 9.5%; Score 86; DB 2; Length 406;
Best Local Similarity 23.2%; Pred. NO. 2.7;
Matches 39; Conservative 27; Mismatches 64; Indels 38; Gaps 6;
QY 6 IFILVIFCAFSSFAQDDSKSTFNLGAGEKFLVYETNKKDSLVPFLNLFGLFGIGSFAQG 65
Db 231 ILLILIFSKYFYWASISSYTYFL--MQRFGLSIQNAQLHLFAFLF-----AVAAG 279
QY 66 DILGSSLILGFDVAGIGLITGAYLDIKDFDNNAKKADFKWTGKGMMLAGVVTMAVTRL 125
Db 280 TVIGGPV---GDKIGRKYVWGSILGVAPFTLILPYASLHWT-----GVLT----- 322
QY 126 TEIVLPFTFANNYNRKLKNSLNIALGGFEPFDFINMQASALGFLSF 173
Db 323 --VIIGFILASAFSAILVYAQLPG-----RIGMVSGLFFGFAP 360

Search completed: March 14, 2003, 14:25:44
Job time : 15.264 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:21:36 ; Search time 7.29214 Seconds
(without alignments)
1006.744 Million cell updates/sec

Title: US-09-508-487-23

Perfect score: 908

Sequence: 1 MNKFLIFILVIFCAFSSFAQ.....DINMQASALGGLSPKKS 177

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|---------------|---------------------|
| 1 | 88.5 | 9.7 | 451 | 1 PTKC_ECOLI | P37189 escherichia |
| 2 | 87 | 9.6 | 1290 | 1 BXB_CLOBO | P10844 clostridium |
| 3 | 86 | 9.5 | 406 | 1 FSR_ECOLI | P52067 escherichia |
| 4 | 85 | 9.4 | 235 | 1 CCDA_BACSU | P45706 bacillus su |
| 5 | 85 | 9.4 | 692 | 1 PHSG_AQUAE | O66932 aquifex aeo |
| 6 | 83.5 | 9.2 | 449 | 1 YK66_YEAST | P36091 saccharomyc |
| 7 | 81 | 8.9 | 174 | 1 NM6M_RABIT | O79438 oryctolagus |
| 8 | 81 | 8.9 | 371 | 1 COXB_AGRIT4 | O59159 agrobacteri |
| 9 | 79.5 | 8.8 | 436 | 1 DCTA_PSEAE | Q914f5 pseudomonas |
| 10 | 79.5 | 8.8 | 933 | 1 SIAP_CAMFE | P35827 campylobact |
| 11 | 78.5 | 8.6 | 935 | 1 EAE_ECOLI1 | O31000 escherichia |
| 12 | 78.5 | 8.6 | 936 | 1 EAE_CITFR | Q07591 citrobacter |
| 13 | 78.5 | 8.6 | 939 | 1 EAE_ECO27 | P19809 escherichia |
| 14 | 78 | 8.6 | 264 | 1 YA96_MYCPN | P75596 mycoplasma |
| 15 | 76.5 | 8.4 | 1103 | 1 YG37_BPARI | Q9g0b5 bacteriophia |
| 16 | 76.5 | 8.4 | 1938 | 1 MYHD_HUMAN | Q9ukx3 homo sapien |
| 17 | 75.5 | 8.3 | 934 | 1 EAE_ECO57 | P43261 escherichia |
| 18 | 75.5 | 8.3 | 1077 | 1 HGPA_HAEIN | Q9za21 haemophilus |
| 19 | 74.5 | 8.2 | 276 | 1 NOSY_PSEST | P19845 pseudomonas |
| 20 | 74 | 8.1 | 359 | 1 OMPA_SERMA | P04845 serratia ma |
| 21 | 74 | 8.1 | 362 | 1 FTBP_ADE08 | P36845 human adeno |
| 22 | 74 | 8.1 | 580 | 1 P69_MYCHR | P15362 mycoplasma |
| 23 | 73.5 | 8.1 | 188 | 1 DCD_HELPRI | Q82kd0 helicobacte |
| 24 | 73.5 | 8.1 | 419 | 1 CMLA_PSEAE | P32482 pseudomonas |
| 25 | 72.5 | 8.0 | 836 | 1 VG26_BPML5 | Q05233 mycobacteri |
| 26 | 72 | 7.9 | 493 | 1 SECY_ARCFU | O28377 archaeglob |
| 27 | 72 | 7.9 | 607 | 1 LPPA_LAGIA | Q9c9i8 lactococcus |
| 28 | 72 | 7.9 | 638 | 1 SGAG_BOVIN | Q28039 bos taurus |
| 29 | 72 | 7.9 | 670 | 1 OATP_HUMAN | P46721 homo sapien |
| 30 | 72 | 7.9 | 1072 | 1 CARB_PSEAE | P38100 pseudomonas |
| 31 | 71.5 | 7.9 | 392 | 1 FXB4_YEAST | Q06205 saccharomyc |
| 32 | 71.5 | 7.9 | 444 | 1 DCTA_RHILE | Q01857 rhizobium l |
| 33 | 71.5 | 7.9 | 853 | 1 PHS1_DICDI | Q00766 dictyosteli |

| | | | | | |
|----|------|-----|-----|--------------|--------------------|
| 34 | 71 | 7.8 | 159 | 1 VT3A_CAPVI | P18387 capripoxvir |
| 35 | 71 | 7.8 | 231 | 1 Y236_CAMJE | Q9piq8 campylobact |
| 36 | 71 | 7.8 | 523 | 1 Y236_RICCO | Q41144 ricinus com |
| 37 | 71 | 7.8 | 612 | 1 PLB1_PENCH | P39457 penicillium |
| 38 | 70.5 | 7.8 | 268 | 1 DCTA_CAMJE | Q46100 campylobact |
| 39 | 70.5 | 7.8 | 393 | 1 NCAP_PVM | P26589 pneumonia v |
| 40 | 70.5 | 7.8 | 581 | 1 FTBP_ADE05 | P11818 human adeno |
| 41 | 70 | 7.7 | 319 | 1 RLA0_ORYSA | P41095 oryza sativ |
| 42 | 70 | 7.7 | 355 | 1 Y816_DROME | Q9vaf0 drosophila |
| 43 | 70 | 7.7 | 371 | 1 COB2_RHIME | P72299 rhizobium m |
| 44 | 70 | 7.7 | 377 | 1 GBB_SOLTU | P93563 solanum tub |
| 45 | 70 | 7.7 | 442 | 1 DTA2_RHILO | Q986r8 rhizobium l |

ALIGNMENTS

| | | | | | |
|----------|---|-----------|------|---------|--|
| RESULT 1 | | | | | |
| ID | PTKC_ECOLI | STANDARD; | PRT; | 451 AA. | |
| AC | P37189; P76411; | | | | |
| DT | 01-OCT-1994 (Rel. 30, Created) | | | | |
| DT | 01-NOV-1997 (Rel. 35, Last sequence update) | | | | |
| DT | 15-JUN-2002 (Rel. 41, Last annotation update) | | | | |
| DE | PTS system, galactitol-specific IIC component (EIIC-GAT) (Galacticol- permease IIC component) (Phosphotransferase enzyme II, C component). | | | | |
| GN | GATC OR B2092 OR Z3255 OR ECS2895. | | | | |
| OS | Escherichia coli, and | | | | |
| OS | Escherichia coli O157:H7. | | | | |
| OC | Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; | | | | |
| OC | Escherichia. | | | | |
| OX | NCBI_TaxID=562, 83334; | | | | |
| RN | [1] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RC | STRAIN=EC3132; | | | | |
| RX | MEDLINE=95290497; PubMed=7772602; | | | | |
| RA | Nobelmann B., Lengeler J.W.; | | | | |
| RT | "Sequence of the gat operon for galactitol utilization from a | | | | |
| RT | wild-type strain EC3132 of Escherichia coli."; | | | | |
| RL | Biochim. Biophys. Acta 1262:69-72(1995). | | | | |
| [2] | | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RN | STRAIN=K12 / MG1655; | | | | |
| RC | MEDLINE=97426617; PubMed=9278503; | | | | |
| RX | Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., | | | | |
| RA | Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., | | | | |
| RA | Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., | | | | |
| RA | Mau B., Shao Y.; | | | | |
| RT | "The complete genome sequence of Escherichia coli K-12."; | | | | |
| RL | Science 277:1453-1474(1997). | | | | |
| [3] | | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RN | STRAIN=K12; | | | | |
| RC | MEDLINE=97251358; PubMed=9097040; | | | | |
| RX | Itoh T., Aiba H., Baba H., Kimura S., Kitakawa M., Kitagawa M., | | | | |
| RA | Isono K., Kasai H., Kimura S., Kitakawa M., Mori T., Motomura K., | | | | |
| RA | Makino K., Miki T., Mizobuchi K., Mori H., Nishio Y., Oshima T., | | | | |
| RA | Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., | | | | |
| RA | Saito N., Sankel G., Seki Y., Sivasubramanian S., Tagami H., | | | | |
| RA | Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.; | | | | |
| RT | "A 460-kb DNA sequence of the Escherichia coli K-12 genome | | | | |
| RT | corresponding to the 40.1-50.0 min region on the linkage map."; | | | | |
| RL | DNA Res. 3:379-392(1996). | | | | |
| [4] | | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RN | STRAIN=O157:H7 / EDL933 / ATCC 700927; | | | | |
| RC | MEDLINE=21074935; PubMed=11206551; | | | | |
| RX | Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D., | | | | |
| RA | Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A., | | | | |
| RA | Posral G., Hackett J., Klink S., Boutin A., Shao Y., Miller L., | | | | |
| RA | Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoukis K., | | | | |
| RA | Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C., | | | | |
| RA | Weich R.A., Blattner F.R.; | | | | |


```
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M81186; AAA23211.1; -.
CC EMBL: Z11934; CAA77991.1; -.
CC EMBL: X70817; CAA50148.1; -.
CC PIR: S07128; S07128.
CC PIR: S07155; S07155.
CC PIR: S08562; S08562.
CC PIR: S08573; S08573.
CC PIR: S08574; S08574.
CC PIR: A48940; A48940.
CC HSP: P10845; 3BTA.
CC MR0P: M27.002; -.
CC InterPro: IPR000395; Bontoxilysin.
CC InterPro: IPR000130; Zn_MTPeptidase.
CC Pfam: PF01742; Peptidase_M27; 1.
CC PRINTS: PR00760; BONTOXILYSIN.
CC ProDom: PD001963; Bontoxilysin; 1.
CC PROSITE: PS00142; ZINC_PROTEASE; 1.
CC Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
CC INIT_MET 0
CC CHAIN 1 440 BOTULINUM NEUROTOXIN B, LIGHT-CHAIN.
CC CHAIN 441 1290 BOTULINUM NEUROTOXIN B, HEAVY-CHAIN.
CC METAL 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
CC ACT_SITE 230 230 BY SIMILARITY.
CC METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
CC DISULFID 436 445 INTERCHAIN (PROBABLE).
CC CONFLICT 29 29 T -> M (IN REF. 4).
CC CONFLICT 217 217 R -> G (IN REF. 2).
CC CONFLICT 224 224 A -> S (IN REF. 2).
CC CONFLICT 463 463 S -> R (IN REF. 4).
CC SEQUENCE 1290 AA; 150670 MW; D21746E2C024DF43 CRC64;
CC -----
Query Match 9.68; Score 87; DB 1; Length 1290;
Best Local Similarity 22.88; Pred. No. 3.6;
Matches 38; Conservative 31; Mismatches 58; Indels 40; Gaps 7;
CC 37 VYETNKKDS-----LVPFL---LNLFLGFGIGSFAQG-DILGSLILGFD-----AV 79
CC 601 VIEANKSTWDKDIADISLIVPYIGIALNVGNETAKGFENAFETAGASILLEFPELLIP 660
CC 80 GIGLILGAYLD-----IKFDNNNAKKADFKWTWKGKMLAGVYVTVAVTRL----- 125
CC 661 VVGAFLESYIDNKKIKTIKTIDNALTNRKNSDMYGLIVAQLSTVNTQFTYKEGMYK 720
CC 126 -----TEVLVPTFANNYRKLKNSLNTALGCFEPSPDINMQA 164
CC 721 ALNYQAQALEEIIKYRY-NIYSEKSNINIDFNDSKLNENGINQA 766
CC -----
RESULT 3
FSR_ECOLI
ID FSR_ECOLI STANDARD; PRT; 406 AA.
AC P52067;
DT 01-OCT-1996 (Rel. 34, Created)
```

```
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fosmidomycin resistance protein.
GN FSR OR B0479.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OC NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / DH5-alpha;
RX MEDLINE=97074653; PubMed=8917080;
RA Fujisaki S., Ohnuma S.-I., Horiuchi T., Takahashi I., Tsukui S.,
RA Nishimura Y., Nishino T., Kitabatake M., Inokuchi H.;
RA "Cloning of a gene from Escherichia coli that confers resistance to
RA fosmidomycin as a consequence of amplification.";
RL Gene 175:83-87(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MGI655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RA Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,
RA Duncan M., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
RA Lew H., Lin D., Namath A., Oefner P., Schramm S., Davis R.W.;
RA Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CONFERS THE RESISTANCE AGAINST FOSMIDOMYCIN.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D73370; BAA1120.1; -.
CC EMBL: AE000154; AAC73581.1; -.
CC EMBL: U82664; BAB40233.1; -.
CC EcoGene: EGI3268; fsr.
CC Antibiotic resistance; Transmembrane; Complete proteome.
CC TRANSMEM 43 63 POTENTIAL.
CC TRANSMEM 64 84 POTENTIAL.
CC TRANSMEM 103 123 POTENTIAL.
CC TRANSMEM 152 172 POTENTIAL.
CC TRANSMEM 178 198 POTENTIAL.
CC TRANSMEM 226 246 POTENTIAL.
CC TRANSMEM 267 287 POTENTIAL.
CC TRANSMEM 295 315 POTENTIAL.
CC TRANSMEM 320 340 POTENTIAL.
CC TRANSMEM 354 374 POTENTIAL.
CC TRANSMEM 379 399 POTENTIAL.
CC SEQUENCE 406 AA; 43305 MW; E1FC88D0E3F41F3 CRC64;
CC -----
Query Match 9.58; Score 86; DB 1; Length 406;
Best Local Similarity 23.28; Pred. No. 1.4;
Matches 39; Conservative 27; Mismatches 64; Indels 38; Gaps 6;
CC 6 IFILVIFCAFSFAQDDSKSTFNLGAGEKFLVYVETNKKDSLVPLFLNLFLGFGIGSFAQG 65
CC 231 ILLILFSKYPYMASISSYTYFL--MQKFGLSIQNAQLHLFAFLF-----AVAAG 279
CC 66 DILGSLILGPDVAVGIGLILTGAYLDIKDFDNNNAKKADFKWTWKGKMLAGVYVTVAVTRL 125
CC 280 TVIGGPV---GDKIGRKYVINGSILGVAPFTLILPYASLHWT-----GVLT----- 322
```



```
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Anjay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358(1998).
CC -1- FUNCTION: PHOSPHORYLASE IS AN IMPORTANT ALLOSTERIC ENZYME IN
CC CARBOHYDRATE METABOLISM. ENZYMES FROM DIFFERENT SOURCES DIFFER IN
CC THEIR REGULATORY MECHANISMS AND IN THEIR NATURAL SUBSTRATES.
CC HOWEVER, ALL KNOWN PHOSPHORYLASES SHARE CATALYTIC AND STRUCTURAL
CC PROPERTIES (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: [(1,4)-alpha-D-glucosyl](N) + phosphate =
CC [(1,4)-alpha-D-glucosyl](N-1) + alpha-D-glucose 1-phosphate.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GLYCOGEN PHOSPHORYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000704; AAC06896.1; -.
CC HSP: P00489; 3AMV.
CC InterPro: IPR000811; GT_35.
CC Pfam: PF00343; phosphorylase; 1.
CC PROSITE: PS00102; PHOSPHORYLASE; 1.
CC Transferrase: Glycosyltransferase; Carbohydrate metabolism;
CC Glycogen metabolism: Pyridoxal phosphate; Complete proteome.
CC BINDING 586 586 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC SEQUENCE 692 AA; 81158 MW; 4DC0EA2B2E58BEA4 CRC64;
CC -----
Query Match 9.4%; Score 85; DB 1; Length 692;
Best Local Similarity 29.9%; Pred. No. 2.9;
Matches 46; Conservative 28; Mismatches 71; Indels 56; Gaps 10;
CC
QY 4 FLIFILVIFAFSFAODSKSTNLGAGEKFLVYETNKKDSLVPFL-----LNLF 54
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 63 FISHFYVSLYKTYMNRHSK-----YEDTYKKPIV-FLSPXYGLHHTLIY 108
QY 55 LGFGIGSFAQGDILGGSLLILGFDVAGVIGLITLGTAYLDIK-----DFDNNAKK----- 101
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 109 AG-GLG-FLAGDILKESDGLGFLGIVGFMYPQGVYKQIRVDGWDQDLDAQNKELMPV 166
QY 102 ---ADFKWTGKGWMLA-----GVVTMAVTRLTEIVLPFTFANN---YNRKLNLSLMA 149
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 167 KKVLDKGKWLKCYVYVDERVYGVGVMEVNGKTKLYLLDFTNVEPTPNREISRL---- 223
QY 150 LGGFEPSPDINMGQASALGFG 170
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 224 ---YVPDKDLRLRQOIVLFG 241
RESULT 6
YK6_YEAST STANDARD; PRT; 449 AA.
AC P36091;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 49.6 kDa protein in ELM1-PR12 intergenic region.
GN YKL046C OR YKL259.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OC NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=94205268; PubMed=8154189;
RA Purnelle B., Tettelin H., van Dyck L., Skala J., Coffeau A.;
RT "The sequence of a 17.5 kb DNA fragment on the left arm of Yeast
RT chromosome XI identifies the protein kinase gene ELM1, the DNA
RT primase gene PRI2, a new gene encoding a putative histone and seven
RT new open reading frames.";
RL Yeast 9:1379-1384(1993).
CC -1- SIMILARITY: TO YEAST YMR238W.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X71621; -. NOT_ANNOTATED_CDS.
CC DR EMBL: Z28046; CAA81881.1; -.
CC PIR: S37867; S37867.
CC SGD: S0001529; YKL046C.
CC InterPro: IPR005198; Glyco_hydro_76.
CC Pfam: PF03663; Glyco_hydro_76; 1.
CC Hypothetical protein; Transmembrane.
CC TRANSMEM 6 26 POTENTIAL.
CC TRANSMEM 431 447 POTENTIAL.
CC SEQUENCE 449 AA; 49565 MW; DECFE9CAFD9579D5 CRC64;
CC -----
Query Match 9.2%; Score 83.5; DB 1; Length 449;
Best Local Similarity 28.8%; Pred. No. 2.5;
Matches 21; Conservative 15; Mismatches 24; Indels 13; Gaps 2;
CC
QY 79 VGIGLILTGAYLDIKD----FDNNAKKADFQWTKGKMLAGVVTMAYTRLTEIVLPFTF 134
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 218 VGVNLISNGTYKYVYDGVSDIDNCTKTSYQWYTNQGLLAG-----SAYLYNFTG 268
QY 135 ANNYNRKLNLSN 147
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 269 SDLWHTRTKEFLN 281
RESULT 7
NU6M_RABIT STANDARD; PRT; 174 AA.
AC O79438;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 6 (EC 1.6.5.3).
GN MTND6 OR ND6 OR NADH6.
OS Oryctolagus cuniculus (Rabbit).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OC NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98317530; PubMed=9653643;
RA Gissi C., Gullberg A., Arnason U.;
RT "The complete mitochondrial DNA sequence of the rabbit, Oryctolagus
RT cuniculus.";
RL Genomics 50:161-169(1998).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
DR EMBL: AJ001588; CAA04858.1; -.
DR InterPro: IPR001457; Oxidored_q3.
DR Pfam: PF00499; oxidored_q3; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 174 AA; 18728 MW; 18740B56661D09E1 CRC64;

Query Match      8.9%; Score 81; DB 1; Length 174;
Best Local Similarity 23.1%; Pred. No. 1.6;
Matches 36; Conservative 24; Mismatches 54; Indels 42; Gaps 7;

QY 47 VPLLNL-----PLGF-----GIGSFAQGDILGSLILGFDVAGIG-----LILTGA 88
DB 4 VFLLSVMFVMGVGFSKSPSPYIGLGLIVSGGV-GCGIVLSFGSGFLGLMFLIYLGG 62
QY 89 YLDIKDFDNNAKKADKFWTKGKMLAG-----VVTMAVTRLTETIVLPTEFAN 136
DB 63 MLVVFGYTAMATEEPEWGSNMILGMFVLGVLMVGLVYVMWSGDGEIVVDFPKMG 122
QY 137 NYNRKLKNSLNIAGGFEPSF--DINMGQASALGFG 170
DB 123 DW-----VVFEGDEVGLIREDSMGVAALYSYG 149

RESULT 8
OOXB_AGR74
ID OAXB_AGR74 STANDARD; PRT; 371 AA.
AC Q59159;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Opine oxidase subunit B (EC 1.-.-) (Octopine oxidase subunit B).
GN OOXB.
OS Agrobacterium tumefaciens (strain Ach5).
OG Plasmid pTiAch5.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176298;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=94321320; PubMed=8045881;
RA Zanker H., Lurz G., Langridge U., Langridge P., Kreusch D.,
RA Schroeder J.;
RT 'Octopine and nopaline oxidases from Ti plasmids of Agrobacterium
RT tumefaciens: molecular analysis, relationship, and functional
RT characterization.'
RL J. Bacteriol. 176.4511-4517(1994).
CC -!- FUNCTION: OXIDATIVE CLEAVAGE OF OCTOPINE INTO L-ARGININE AND
CC PYRUVATE (BY SIMILARITY).
CC -!- PATHWAY: CATABOLIC UTILIZATION OF OCTOPINE.
CC -!- SUBUNIT: HETERODIMER OF A SUBUNIT AND A B SUBUNIT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z30328; CAA82987.1; -.
CC InterPro: IPR000927; D_aa_oxidase.
CC InterPro: IPR000205; NAD_binding.
CC Pfam: PF01266; DAO; 1.
KW Oxidoreductase; Plasmid.
SQ SEQUENCE 371 AA; 39939 MW; 948705136822AC52 CRC64;

Query Match      8.9%; Score 81; DB 1; Length 371;
Best Local Similarity 27.3%; Pred. No. 3.5;
Matches 38; Conservative 18; Mismatches 43; Indels 40; Gaps 7;

QY 67 ILGGSILIGFDVAGIG-----LILTGAYLDIKDFDNNAKADFK--WTWKGKMLAGVV 118
DB 8 IIGGGLVGASIANGLARSGTKPLVLDGADLDR-----ASRANFALVWVGKGLHAPHYA 62
```

```
QY 119 -----TMAVTRLTETIVL-----PFTFANN-----YNRKLKNSLNIAGGFE 154
DB 63 LWSDSASRRWPTWANTLLDDSGIDVGLQDGAFTFALSEELEARNQDMESIELETNGRA 122
QY 155 PSFDINMGQAS-----ALGFG 170
DB 123 PQFEVLDRQOTLDRVLGIG 141

RESULT 9
DCTA_PSEAE
ID DCTA_PSEAE STANDARD; PRT; 436 AA.
AC Q914F5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE C4-dicarboxylate transport protein.
GN DCTA OR PA1183.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
SEQUENCE FROM N.A.
RX STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.A., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
CC -!- FUNCTION: Responsible for the transport of dicarboxylates such as
CC succinate, fumarate, and malate from the periplasm across the
CC inner membrane (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SODIUM-DICARBOXYLATE SYMPORTER FAMILY
CC (SDF, TC 2.A.23).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE004548; AG04572.1; -.
CC InterPro: IPR001991; Na/dico_symp.
CC Pfam: PF00375; SDF; 1.
CC PRINTS: PR00173; EDTNSPOT.
CC PROSITE: PS00713; NA_DICARBOXYL_SYMP_1; FALSE_NEG.
CC PROSITE: PS00714; NA_DICARBOXYL_SYMP_2; 1.
KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
KW Complete proteome.
FT TRANSMEM 9 28 POTENTIAL.
FT TRANSMEM 48 65 POTENTIAL.
FT TRANSMEM 78 100 POTENTIAL.
FT TRANSMEM 141 163 POTENTIAL.
FT TRANSMEM 189 211 POTENTIAL.
FT TRANSMEM 226 248 POTENTIAL.
FT TRANSMEM 290 312 POTENTIAL.
FT TRANSMEM 327 349 POTENTIAL.
FT TRANSMEM 354 376 POTENTIAL.
SQ SEQUENCE 436 AA; 45993 MW; DCA564DB31A5AC07 CRC64;

Query Match      8.8%; Score 79.5; DB 1; Length 436;
Best Local Similarity 37.1%; Pred. No. 5.6;
```

```
Matches 26; Conservative 7; Mismatches 24; Indels 13; Gaps 3;
QY 21 DSKSTENLGAGEKFLVYETNKKDSLVPFLNLFLGFGIGSFAQGDILGSLILGFDVAV- 79
Db 114 DTSKIAAYAAAGE-----KQSTVDFLMNVPIGVVGAFANGDIL---QVLFFSVLFP 161
QY 80 GIGLILTGAY 89
Db 162 GYALHRLGSY 171

RESULT 10
SLAP_CAMFE
ID SLAP_CAMFE STANDARD; PRT; 933 AA.
AC P35827;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE S-layer protein (Surface array protein) (SAP).
GN SAPA.
OS Campylobacter fetus.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=196;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.
RC STRAIN=84-32 / 23D;
RX MEDLINE=90354448; PubMed=2387868;
RA Blaser M.J., Gotschlich E.C.;
RT "Surface array protein of Campylobacter fetus. Cloning and gene
structure."
RL J. Biol. Chem. 265:14529-14535(1990).
RN [2]
RP ERRATUM.
RX MEDLINE=91035477; PubMed=2229082;
RA Blaser M.J., Gotschlich E.C.;
RL J. Biol. Chem. 265:19372-19372(1990).
CC -1- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA. THIS PROTEIN IS
CRITICAL FOR VIRULENCE.
CC -1- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
S-LAYER WITH HEXAGONAL SYMMETRY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC -----
DB EMBL; J05577; AAA23032.1; -.
KW Cell wall; S-layer.
SQ SEQUENCE 933 AA; 96757 MW; F8C729B4BA5B1E9 CRC64;

Query Match 8.8%; Score 79.5; DB 1; Length 933;
Best Local Similarity 24.4%; Pred. No. 12;
Matches 33; Conservative 19; Mismatches 46; Indels 37; Gaps 5;

QY 73 ILGFDVAVGIG-----LILTAYLDIKDFDNN-----AKKADFKWTWGRGMM 113
Db 652 VINFTADAGSVKLITVKLNDVTALMIVVKIVLDAAAKDTNIALGTAAD-----KALV 705
QY 114 L-AGVVTMAYTRLTETVLPFTFANNYRKLKNSLNIALGGPEPSF-----DINM 161
Db 706 IDTGIETLNITSLVKATSPETTANTVNAKLTDTVTSIIIDGMQITLGHAGTAGTDYSKYSM 765
QY 162 GQASALGGLSEKKS 176
Db 766 IDASALKAGLTFDAS 780

RESULT 11
```

```
EAE_ECOLI
ID EAE_ECOLI STANDARD; PRT; 935 AA.
AC Q31000;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Intimin (Attaching and effacing protein) (Eae protein).
GN EAE OR EAEA.
OS Escherichia coli O111:H-.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=168927;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98187918; PubMed=9529069;
RA Voss E., Paton A.W., Manning P.A., Paton J.C.;
RT "Molecular analysis of Shiga toxinogenic Escherichia coli O111:H-
proteins which react with sera from patients with hemolytic-uremic
syndrome."
RL Infect. Immun. 66:1467-1472(1998).
CC -1- FUNCTION: NECESSARY FOR THE PRODUCTION OF ATTACHING AND EFFACING
LESIONS ON TISSUE CULTURE CELLS. BELIEVED TO MEDIATE ADHERENCE.
CC -1- SUBCELLULAR LOCATION: OUTER SURFACE.
CC -1- SIMILARITY: BELONGS TO THE INTIMIN/INVASIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC -----
DB EMBL; AF025311; AAC69247.1; -.
DR InterPro; IPR003344; Big_1.
DR InterPro; IPR003343; Big_2.
DR InterPro; IPR003535; Intimin.
DR InterPro; IPR002482; LysM.
DR Pfam; PF01476; LysM; 1.
DR Pfam; PF02368; Big_2; 1.
DR Pfam; PF02369; Big_1; 2.
DR PRINTS; PR01369; INTIMIN.
DR SMART; SM00257; LysM; 1.
DR Outer membrane; Virulence.
KW REPEAT
FT REPEAT 65 113 LYSM.
SQ SEQUENCE 935 AA; 101570 MW; 406E79CDC07DEB11 CRC64;

Query Match 8.6%; Score 78.5; DB 1; Length 935;
Best Local Similarity 21.7%; Pred. No. 15;
Matches 34; Conservative 23; Mismatches 45; Indels 55; Gaps 7;

QY 22 DSKSTFNLGAGEKFLVYETNKKDSLVPFLNLFLGFGIGSFAQGDILGSLILGFDVAVGI 81
Db 252 DSRETFANLGAGORE-----FLPENMLGYNV--FIDQDFSGDNLRLG----- 290
QY 82 GLIITGAYLDIKDFDNNAKKADFKWTWKGCMMLAGVVTMAYTRLTETVLPFTFANNYR 140
Db 291 ---IGGETW--RDYFKSVNGYFRMSGWHESYN-----KKDYDE 324
QY 141 KLRKNSLNIALGGPEPSFDINMGQASALGFLSKFSY 177
Db 325 RPANGFDIRNGYLPSPY-----PALGAKLMYEQYY 354

RESULT 12
EAE_CITFR
ID EAE_CITFR STANDARD; PRT; 936 AA.
AC Q07591;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Intimin (Attaching and effacing protein) (Eae protein).
```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:22:51 ; Search time 24.1966 Seconds
(without alignments)
1507.249 Million cell updates/sec

Title: US-09-508-487-23
 Perfect score: 908
 Sequence: 1 MNKFLIFLIVIFACSSFAO.....DINMGOASALGGLSEKSY 177

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```
Database :
SPTRMBL_21:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp_rvirus:*
16: sp.bacteriap:*
17: sp.archaea:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query | | Length | DB | ID | Description |
|------------|-------|-------|-----|--------|--------|--------|--------------|
| | | Match | % | | | | |
| 1 | 908 | 100.0 | 177 | 2 | Q9X3V0 | Q9X3V0 | borrellia ga |
| 2 | 786.5 | 86.6 | 178 | 2 | Q9X3U9 | Q9X3U9 | borrellia af |
| 3 | 737 | 81.2 | 179 | 16 | O51065 | O51065 | borrellia bu |
| 4 | 309 | 33.0 | 161 | 16 | O50896 | O50896 | borrellia bu |
| 5 | 301.5 | 34.2 | 170 | 16 | O50885 | O50885 | borrellia bu |
| 6 | 242.5 | 26.7 | 161 | 2 | Q9S011 | Q9S011 | borrellia bu |
| 7 | 240 | 26.4 | 190 | 16 | O50696 | O50696 | borrellia bu |
| 8 | 170.5 | 18.8 | 95 | 2 | O31325 | O31325 | borrellia bu |
| 9 | 95.5 | 10.5 | 289 | 16 | Q9KXK6 | Q9KXK6 | streptomyce |
| 10 | 93 | 10.2 | 779 | 10 | Q92FB1 | Q92FB1 | listeria in |
| 11 | 91 | 10.0 | 375 | 16 | Q9LHX2 | Q9LHX2 | oryza sativ |
| 12 | 90.5 | 10.0 | 427 | 2 | Q93RE6 | Q93RE6 | vibrio para |
| 13 | 88.5 | 9.7 | 376 | 16 | Q9AWN1 | Q9AWN1 | rhizobium m |
| 14 | 86.5 | 9.5 | 389 | 5 | Q22972 | Q22972 | caenorhabdi |
| 15 | 86.5 | 9.5 | 427 | 2 | Q8VS09 | Q8VS09 | klebsiella |
| 16 | 86.5 | 9.5 | 447 | 16 | O8RFA1 | O8RFA1 | fusobacteri |

ALIGNMENTS

RESULT 1

RESOLUTION 1
Q9X3V0
ID Q9X3V0
PRELIMINARY:
PRT: 177 AA.

Q9A3V0; AC 01-NOV-1999 (TrEMBLrel. 12, Created)
 DTD 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DTD 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Membrane protein P13.

GN Borrelia garinii.
 OS Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=29519;
 RN [1]
 RP SEQUENCE FROM N.A.

| | | |
|-----|---|--|
| RC | STRAIN:1190; | |
| RC | MEDLINE:21189251; PubMed:11292755; | |
| RR | Noppa L., Ostberg Y., Lavrinovitcha M., Bergstrom S.; | |
| RR | "p13, an integral membrane protein of Borrelia burgdorferi, is C- | |
| RR | terminally processed and contains surface-exposed domains."; | |
| RR | Infect. Immun. 69:3323-3334(2001). | |
| RR | EMBL: AF085741; AAD28362.1; - | |
| RR | SEQUENCE 177 AA: 19308 MW: 429928RC6426E1111 CRC64: | |
| DSO | | |

Query Match 100.0%; Score 908; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.6e-73;

0y 1 MNKEI.JETIIVIECAFSSEAODDSKSTENIGAGEKFLVYETNKKDSI.VPFLNLFLGFGIG 60

1 MNKELEIFILVIECAFESFAODDSKSTENIGAGEKEIVVETNKKDSI.VPEI.INI.EI.GEGIG 60

Q: 61 CERNOCYCCST YI CERNAYCICST YI ECAYI NTVDCENNAYKZADENYJTCYCYCWMY KCIWJTM 120

[illegible]

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

A vertical ruler with markings from 0 to 10 cm. The markings are in millimeters, with major ticks every centimeter and minor ticks every millimeter. The ruler is oriented vertically with the 0 mark at the top and the 10 cm mark at the bottom.


```
RESULT 5
O50885 ID O50885 PRELIMINARY; PRT; 170 AA.
DR EMBL; AE001584; AAF07707.1; -.
KW Hypothetical protein; plasmid.
SQ SEQUENCE 161 AA; 17016 MW; 14D4906CC8107CD2 CRC64;

Query Match 26.7%; Score 242.5; DB 2; Length 161;
Best Local Similarity 39.3%; Pred. No. 3.9e-14;
Matches 59; Conservative 25; Mismatches 49; Indels 17; Gaps 6;

OY 18 FAQDSSKSTFNLAGAE--KELVYETNKKDSLVPFLNLFGLGFGISFAQGDILGGSLLIG 75
Db 4 FAQEKLEK----GVGDIATVMKYESKATILAPLLNLLSLGIGSFVQGDYIGGCALLG 59
OY 76 FDAVGIGLILTCAYL--DIKDFDNNNAKADFKWTGKGMMLAGV--VTMAVTRTEIVLP 131
Db 60 SOVLG-GILIMAGYMTGDIGFVTESTA-----TVITGVLSGIGLTAASYYITGIIP 112
OY 132 FTFANNYNRKLNLSNIALGGFEPFSFDINM 161
Db 113 KFEANRYNADLKKRLGIALAGLEPNFDIGI 142

RESULT 7
O50696 ID O50696 PRELIMINARY; PRT; 190 AA.
DR EMBL; AE000784; AAC65989.1; -.
KW TIGR; BBH41; -.
SQ SEQUENCE 190 AA; 21813 MW; A80FB8BCAD4B13EB CRC64;

Query Match 26.4%; Score 240; DB 16; Length 190;
Best Local Similarity 41.3%; Pred. No. 7.9e-14;
Matches 59; Conservative 22; Mismatches 50; Indels 12; Gaps 5;

OY 36 LVYETNKK-DSLVPFLNLFGLGFGISFAQGDILGGSILGSDAVGIGLILTCAYLIDIKD 94
Db 59 LQYEKEQLIDLWIPVNLNLFPLFGVGSFVQGDYIGGCGTGLGFLNLLGLTLTTGV-IQMK- 116
OY 95 FDNNAKADFKWTGCKGMWLAGVVTMAVTRTEIVLPFFANNYNRKLNLSNIALGGFE 154
Db 117 ---NLKKEPASIS-SMILLSGMLTFGSSYLISIVLPVLFEDRYKLNLMRIIDELAGFE 172
OY 155 PSFDINMGQASALGFLGFKKSY 177
Db 173 PNLDIGMN-----GFQLSFKKSY 190

RESULT 8
O31325 ID O31325 PRELIMINARY; PRT; 161 AA.
DR EMBL; AE000784; AAC65989.1; -.
KW TIGR; BBH41; -.
SQ SEQUENCE 161 AA; 17016 MW; 14D4906CC8107CD2 CRC64;

Query Match 33.2%; Score 301.5; DB 16; Length 170;
Best Local Similarity 40.1%; Pred. No. 2.2e-19;
Matches 71; Conservative 23; Mismatches 70; Indels 13; Gaps 3;

OY 1 MNKILFLIFVIFCAFSSPAODSKSTFNLAGAEKFLVYETNKKDSLVPFLNLFGLGIG 60
Db 7 MKKILTLILSLTIQIPATOD-KLEKSVGSIETIMKYSEKATILAPLLNLFGLGIG 65
OY 61 SFAQGDILGSLILGFDVAVGIGLILTCAYLIDIKDFDNNNAKADFKWTGKGMMLAGVVTM 120
Db 66 SFVQGDYIGGAVLGSQSLGILGILTCAGNLTGTDDETRA-----TTGHIITIGVCTI 118
OY 121 AVTRTEIVLPFFANNYNRKLNLSNIALGGFEPFSFDINMGQASALGFLGFKKSY 177
Db 119 IASHIASLIIPFTFANKHNANLKKRLGIDGAGFEPNFDIGIS-----GFQLSFKKRY 170

RESULT 6
O9S011 ID O9S011 PRELIMINARY; PRT; 161 AA.
DR EMBL; AE000789; AAC66194.1; -.
KW TIGR; BBH31; -.
SQ SEQUENCE 170 AA; 18227 MW; OE71B633EA80FB22 CRC64;

Query Match 33.2%; Score 301.5; DB 16; Length 170;
Best Local Similarity 40.1%; Pred. No. 2.2e-19;
Matches 71; Conservative 23; Mismatches 70; Indels 13; Gaps 3;

OY 1 MNKILFLIFVIFCAFSSPAODSKSTFNLAGAEKFLVYETNKKDSLVPFLNLFGLGIG 60
Db 7 MKKILTLILSLTIQIPATOD-KLEKSVGSIETIMKYSEKATILAPLLNLFGLGIG 65
OY 61 SFAQGDILGSLILGFDVAVGIGLILTCAYLIDIKDFDNNNAKADFKWTGKGMMLAGVVTM 120
Db 66 SFVQGDYIGGAVLGSQSLGILGILTCAGNLTGTDDETRA-----TTGHIITIGVCTI 118
OY 121 AVTRTEIVLPFFANNYNRKLNLSNIALGGFEPFSFDINMGQASALGFLGFKKSY 177
Db 119 IASHIASLIIPFTFANKHNANLKKRLGIDGAGFEPNFDIGIS-----GFQLSFKKRY 170

RESULT 6
O9S011 ID O9S011 PRELIMINARY; PRT; 161 AA.
DR EMBL; AE000789; AAC66194.1; -.
KW TIGR; BBH31; -.
SQ SEQUENCE 170 AA; 18227 MW; OE71B633EA80FB22 CRC64;

Query Match 33.2%; Score 301.5; DB 16; Length 170;
Best Local Similarity 40.1%; Pred. No. 2.2e-19;
Matches 71; Conservative 23; Mismatches 70; Indels 13; Gaps 3;

OY 1 MNKILFLIFVIFCAFSSPAODSKSTFNLAGAEKFLVYETNKKDSLVPFLNLFGLGIG 60
Db 7 MKKILTLILSLTIQIPATOD-KLEKSVGSIETIMKYSEKATILAPLLNLFGLGIG 65
OY 61 SFAQGDILGSLILGFDVAVGIGLILTCAYLIDIKDFDNNNAKADFKWTGKGMMLAGVVTM 120
Db 66 SFVQGDYIGGAVLGSQSLGILGILTCAGNLTGTDDETRA-----TTGHIITIGVCTI 118
OY 121 AVTRTEIVLPFFANNYNRKLNLSNIALGGFEPFSFDINMGQASALGFLGFKKSY 177
Db 119 IASHIASLIIPFTFANKHNANLKKRLGIDGAGFEPNFDIGIS-----GFQLSFKKRY 170
```

```
ID O31325 PRELIMINARY; PRT; 95 AA.
AC O31325;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE (clone 6) s3.
GN S3.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N40;
RX MEDLINE=96305348; PubMed=8688460;
RA Feng S., Das S., Barthold S.W., Fikrig E.;
RT "Characterization of two genes, p11 and p5, on the Borrelia
RL burgdorferi 49-kilo base linear plasmid.";
RL Blochim. Biophys. Acta 1307:270-272(1996).
DR EMBL; L41151; AAB62281.1; -.
SQ SEQUENCE 95 AA; 10739 MW; D12FC791F8920C3A CRC64;

Query Match 18.8%; Score 170.5; DB 2; Length 95;
Best Local Similarity 45.0%; Pred. No. 5.6e-08;
Matches 36; Conservative 11; Mismatches 32; Indels 1; Gaps 1;

QY 1 MKKFLIFILVFCFAFSFAQDDSKSTFNLGAGEKFLVYETNKKDSLVPFLNLFGLFGIG 60
DB 1 MKKFTLLILFLGLTIEIFATKQNTQRIKGI-ESFNKYDKKKNPGIPFLNLFGLFGIG 59

QY 61 SFAQGDILGSLILGPDVAG 80
DB 60 SFVQGDYIGGSGVLGNLLG 79

RESULT 9
Q9XXK6
ID Q9XXK6 PRELIMINARY; PRT; 289 AA.
AC Q9XXK6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative integral membrane protein.
GN SCO2214 OR SCC53.05.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL357591; CAB93434.1; -.
SQ SEQUENCE 289 AA; 29123 MW; 89AD68D9D805BE10 CRC64;

Query Match 10.5%; Score 95.5; DB 16; Length 289;
Best Local Similarity 28.8%; Pred. No. 1;
Matches 36; Conservative 13; Mismatches 43; Indels 33; Gaps 5;

QY 71 SLILGPDVAGIG---LILTGAYLIDIKDFDNNAKADFKTWKGKMLAGVVTWAVTRLTE 127
DB 5 SLTRG-DGVVIGRAVLLIIASFLDIYSIDGAPDSADIPSLWGSVVLGVLAGVIGAL 63
```

```
QY 128 IVLPFTFANNYNRKLNKSLNIA---LGGFEPFSFDI-----NMGOASA 166
DB 64 VVV-----NRAMPQPKKVGLDLQFGIATVFAAWCALGNIFDPAGAFDNDGASG 115

QY 167 LGFGL 171
DB 116 IGAGM 120

RESULT 10
Q92FB1
ID Q92FB1 PRELIMINARY; PRT; 779 AA.
AC Q92FB1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein lin0195.
GN Lin0195.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OX Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / SEROVAR 6A;
RX PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kref J., Kuhn M., Kunst F., Kurapkhat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlend J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL596163; CAC95428.1; -.
DR ListList; LIN00195; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 779 AA; 89623 MW; AB7B7D066C715261 CRC64;

Query Match 10.2%; Score 93; DB 16; Length 779;
Best Local Similarity 26.6%; Pred. No. 5.7;
Matches 45; Conservative 24; Mismatches 52; Indels 48; Gaps 9;

QY 14 AFSSFAQ-----DSSKSTFNLGAGEKFL-VYETNKKDSLVPFLNLFGLFGIGSPAQ 65
DB 614 AFELSEKYPEVRAQKQSPFMNEAREEFLAFAFETGRTFVGFV---LG---GVFSEG 667

QY 66 DILGSSILILGFDVAGIGLITGAYLD-IKDFDNNAKADFKTWKGKMLA-----GVV 118
DB 668 VDLRGERLICTAIVGGLAQMNVESDLIKDYNE-----TIGRGFDYAYIPGMKV 719

QY 119 TMAVTRLTE-----IVLPFTFAN-----NYNRKLKSLNLI 148
DB 720 LQAVGRVIRCEKDGQGVVLLIEERFATERYKALFPAHNNHARTVKNLTQI 768

RESULT 11
Q9LHX2
ID Q9LHX2 PRELIMINARY; PRT; 375 AA.
AC Q9LHX2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Similar to Arabidopsis thaliana chromosome II sequence from clones
DE T22013 (OSUNBa0036717.30 protein).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
```

```
OC Ehrhartoidea; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
  clone:PO515G01.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
  clone:OSJNBa0036J17.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP001633; BAA94232.1; -.
DR EMBL; AP003104; BAB55738.1; -.
DR InterPro; IPR001087; Lipase_GDSL.
SQ SEQUENCE 375 AA; 40234 MW; 238D34ACAD934E93 CRC64;

Query Match 10.08; Score 91; DB 10; Length 375;
Best Local Similarity 24.38; Pred. No. 3.6;
Matches 42; Conservative 26; Mismatches 61; Indels 44; Gaps 8;

Qy 1 MNKFLIFILVIFCAFSSFAODSKSTFNIG-----AGEKFLVYETNKKDSLIV---PFL 51
Db 1 MQKVLWQIQLVLLSFSFSEVETDVASIFSGDSFDGNIYLYGPARTDLVMTKPPYGM 60
Qy 52 NLFGLFGTGAQGDILGGSLILGFDVAVIGL-ILTGAYLDIKDFDNNAKKADFKWTWKG 110
Db 61 TFF-----DHPGRLSDGRLIIDFIAEALGLPLPPSFAANRSFEHCA----- 103
Qy 111 GMMLAGVVTMAVTRLTETVLFTFANNYNRKLNLSNLTALG-----GFEP 156
Db 104 -----NFATAGCTALDRAFFVANNFT--VMSPFNISLQDLGLWLDGMKPS 146

RESULT 12
Q93RE6 Query Match 9.78; Score 88.5; DB 16; Length 376;
ID Best Local Similarity 26.18; Pred. No. 6;
AC Q93RE6 PRELIMINARY; PRT; 427 AA.
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Adhesin.
GN VAG.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RA Akeda Y., Honda T.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047560; BAB59008.1; -.
DR InterPro; IPR02528; Mate.
DR Pfam; PF01554; UPF0013.2.
DR TIGRFAMs; TIGR00797; mate; 1.
SQ SEQUENCE 427 AA; 45909 MW; 0B608B9DC70E3C71 CRC64;

Query Match 10.08; Score 90.5; DB 2; Length 427;
Best Local Similarity 26.18; Pred. No. 4.6;
Matches 52; Conservative 23; Mismatches 65; Indels 59; Gaps 11;

Qy 5 LIFILVIFCAFSSFAQ-----DDSKSTFNILGAGEKFLVYETNKKDSLIV----- 47
Db 73 MVFAALAVCLFVFPQPIGLMTDSQEVIELGS--SYLVISSASMFACVASMVGLRA 130
Qy 48 ---PFLNLFGLFGIGS--FAQGDILGSLILGFDVAVIGLILTG-----Y 89
Db 131 MHQPGCLSTFFSGIGLSNVLNWLIFGH--LGFPALGITGAATATVISGAIEVGCLFGY 188

Qy 90 LDIKDFDNNAKKADFKWTWKGMMLAGVYVTVAVTRLTETVLFTFANNYNRKLNLSNLTAL 149
Db 189 LWL-----KKHIAFGWGD--IRASLVLDKITRFLSLSLPTTF-----NFLAWA 230
Qy 150 LGGFEPSPFDINWQASALG 168
Db 231 GGLF--AYHAIMGAGVQVG 247

RESULT 13
Q92WNI Query Match 9.78; Score 88.5; DB 16; Length 376;
ID Best Local Similarity 26.18; Pred. No. 6;
AC Q92WNI PRELIMINARY; PRT; 376 AA.
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Putative sugar ABC transporter permease protein.
GN RB0301 OR SMB20315.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RA MEDLINE-21396508; PubMed-11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P., Gouzy J.,
  Vorhoelster F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
  Ra Golding B., Puehler A.;
RT "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
  fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL; AL603643; CAC48701.1; -.
DR InterPro; IPR001851; Bac_inmem_transp.
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 376 AA; 39929 MW; 7DB5E8189CFAF827 CRC64;

Query Match 9.78; Score 88.5; DB 16; Length 376;
Best Local Similarity 26.18; Pred. No. 6;
Matches 49; Conservative 19; Mismatches 75; Indels 45; Gaps 10;

Qy 2 NKFLIFILVIFCAFSSFAQDDSKSTFNILGAGEKFLVYETNKKDSLIVPFLNLFGLFGIG 60
Db 65 NVLVVALAVF-----AGETFLSLNFQSMSAQVPELALLALGVMLA 106
Qy 61 SFAQGDILGSLILGFDVAVIGLIL---TGAYLDIKDFDNNAKKAD--FKWTWKGMMMLA 115
Db 107 MIAGG-----GIDLSGIALANLAGVGSYLLVRDW--VSADAPLAFSMLFAAMALLI 157
Qy 116 GV-----VTMAVTRLTETVLFTFANNYNRKLNLSNLTALGGEPEPFDINWQAS 165
Db 158 GLAGLLNGALIAFAGLPIIATIGTQLFTGLAVAFNGSAITLGYIEP-LD-NFGWTP 215
Qy 166 ALGFGLSF 173
Db 216 VLGVPMCF 223

RESULT 14
Q22972 Query Match 9.78; Score 88.5; DB 16; Length 376;
ID Best Local Similarity 26.18; Pred. No. 6;
AC Q22972 PRELIMINARY; PRT; 389 AA.
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE F28A12.4 protein.
GN F28A12.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
  Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:23:06 ; Search time 10.9382 Seconds
(without alignments)
476.116 Million cell updates/sec

Title: US-09-508-487-23
Perfect score: 908
Sequence: 1 MNRFLPILVIFCAFSSFAQ.....DINMGQASALGFLGFSKKSX 177

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|----|---------------------|
| 1 | 87 | 9.6 | 858 | 4 | US-09-255-829-22 |
| 2 | 87 | 9.6 | 858 | 4 | US-09-255-829-29 |
| 3 | 87 | 9.6 | 1169 | 4 | US-09-255-829-20 |
| 4 | 77 | 8.5 | 243 | 4 | US-09-134-001C-2961 |
| 5 | 74.5 | 8.2 | 934 | 4 | US-08-840-466A-19 |
| 6 | 74.5 | 8.2 | 934 | 4 | US-09-696-188B-19 |
| 7 | 71 | 7.8 | 612 | 3 | US-09-295-186-16 |
| 8 | 70.5 | 7.8 | 95 | 4 | US-08-851-362D-37 |
| 9 | 70.5 | 7.8 | 116 | 4 | US-08-851-362D-49 |
| 10 | 70.5 | 7.8 | 344 | 4 | US-09-134-001C-4175 |
| 11 | 70.5 | 7.8 | 581 | 4 | US-09-393-627B-1 |
| 12 | 69 | 7.6 | 376 | 1 | US-08-608-241-2 |
| 13 | 69 | 7.6 | 376 | 2 | US-08-922-182-2 |
| 14 | 69 | 7.6 | 376 | 2 | US-09-919-953-2 |
| 15 | 69 | 7.6 | 376 | 4 | US-09-192-983-2 |
| 16 | 69 | 7.6 | 415 | 4 | US-09-134-001C-5101 |
| 17 | 68 | 7.5 | 165 | 4 | US-08-858-207A-520 |
| 18 | 68 | 7.5 | 539 | 2 | US-08-467-963C-2 |
| 19 | 68 | 7.5 | 539 | 2 | US-08-838-189D-2 |
| 20 | 68 | 7.5 | 539 | 3 | US-08-852-344D-2 |
| 21 | 68 | 7.5 | 539 | 3 | US-08-344-639E-2 |
| 22 | 68 | 7.5 | 539 | 4 | US-08-467-969A-2 |
| 23 | 68 | 7.5 | 539 | 4 | US-08-467-961A-2 |
| 24 | 68 | 7.5 | 539 | 4 | US-08-001-554A-2 |
| 25 | 68 | 7.5 | 691 | 4 | US-09-620-412C-313 |
| 26 | 67.5 | 7.4 | 349 | 4 | US-09-134-001C-4004 |
| 27 | 67.5 | 7.4 | 538 | 4 | US-09-134-001C-4633 |

| | | | | | | |
|----|------|-----|------|---|---------------------|-------------------|
| 28 | 67.5 | 7.4 | 1861 | 2 | US-08-790-912-4 | Sequence 4, Appl |
| 29 | 67 | 7.4 | 239 | 2 | US-08-933-750C-22 | Sequence 22, Appl |
| 30 | 67 | 7.4 | 239 | 4 | US-09-234-613-32 | Sequence 22, Appl |
| 31 | 67 | 7.4 | 337 | 1 | US-08-312-387B-5 | Sequence 5, Appl |
| 32 | 67 | 7.4 | 337 | 1 | US-08-312-387B-12 | Sequence 12, Appl |
| 33 | 67 | 7.4 | 337 | 1 | US-08-683-426-5 | Sequence 5, Appl |
| 34 | 67 | 7.4 | 337 | 1 | US-08-683-426-12 | Sequence 12, Appl |
| 35 | 67 | 7.4 | 337 | 1 | US-08-683-458-5 | Sequence 5, Appl |
| 36 | 67 | 7.4 | 337 | 1 | US-08-683-458-12 | Sequence 12, Appl |
| 37 | 67 | 7.4 | 337 | 2 | US-08-878-360-5 | Sequence 5, Appl |
| 38 | 67 | 7.4 | 337 | 2 | US-08-878-360-12 | Sequence 12, Appl |
| 39 | 67 | 7.4 | 337 | 3 | US-08-478-140B-5 | Sequence 5, Appl |
| 40 | 67 | 7.4 | 337 | 4 | US-09-333-412-5 | Sequence 5, Appl |
| 41 | 67 | 7.4 | 337 | 4 | US-09-333-412-12 | Sequence 12, Appl |
| 42 | 67 | 7.4 | 337 | 4 | US-09-338-943-5 | Sequence 5, Appl |
| 43 | 67 | 7.4 | 346 | 4 | US-09-134-001C-5196 | Sequence 5196, Ap |
| 44 | 67 | 7.4 | 412 | 4 | US-09-423-439-10 | Sequence 10, Appl |
| 45 | 67 | 7.4 | 639 | 4 | US-09-134-001C-5661 | Sequence 5661, Ap |

ALIGNMENTS

RESULT 1
US-09-255-829-22
; Sequence 22, Application US/09255829
; Patent No. 6461617
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN, & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/255.829
; FILING DATE: 23-FEB-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB97/02273
; FILING DATE: 22-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/782.893
; FILING DATE: 27-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1581.0130002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 858 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-255-829-22

Query Match 9.6%; Score 87; DB 4; Length 858;
Best Local Similarity 22.8%; Pred. NO. 0.22;
Matches 38; Conservative 31; Mismatches 59; Indels 40; Gaps 7;

```

QY   37 VYETNKKDS-----LYPFL---LNLPLGFGIGSFAQG-DILGGSLILGFD-----AV 79
      ! | | : : ||| : ||| : | | : | | : | | : | | : | | : | | : | |
Db    602 VIANKSNWMDKIADISLIIVPYIGLALNVGNETA KGFNAFETAGASILLFIPELLIP 661
      : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
QY   80 GIGLILTGA YLD-----IKDFDNNAKKADFKEWTGWKGMMLAGVVTMAVTRL----- 125
      : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
Db    662 VVGAFLESIDYDNKNKIITKDIALTKRNEKWSDMYGLIVAQLWLSVTNTQFYTIKEGMKY 721
      : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
QY   126 -----TEIVLPFTFANNYRKLNKS LNIALGGFEPSFDINMQA 164
      : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
Db    722 ALNYQAQALEI IKRY-NIIYSEKEKSNIIDENDINSKLINESIQNA 767
      : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |

RESULT 2
US-09-255-829-29
; Sequence 29, Application US/09255829
; Patent No. 6461617
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN, & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/255,829
; FILING DATE: 23-FEB-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB97/02273
; FILING DATE: 22-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/782,893
; FILING DATE: 27-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1581.0130002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 858 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-255-829-29

Query Match          9.6%; Score 87; DB 4; Length 858;
Best Local Similarity 22.8%; Pred. No. 0.22;
Matches 38; Conservative 31; Mismatches 58; Indels 40; Gaps

QY   37 VYETNKKDS-----LYPFL---LNLPLGFGIGSFAQG-DILGGSLILGFD-----AV 79
      ! | | : : ||| : ||| : | | : | | : | | : | | : | | : | | : | |
Db    602 VIANKSNWMDKIADISLIIVPYIGLALNVGNETA KGFNAFETAGASILLFIPELLIP 661
      : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
QY   80 GIGLILTGA YLD-----IKDFDNNAKKADFKEWTGWKGMMLAGVVTMAVTRL----- 125
      : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
Db    662 VVGAFLESIDYDNKNKIITKDIALTKRNEKWSDMYGLIVAQLWLSVTNTQFYTIKEGMKY 721
      : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
QY   126 -----TEIVLPFTFANNYRKLNKS LNIALGGFEPSFDINMQA 164
      : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
Db    722 ALNYQAQALEI IKRY-NIIYSEKEKSNIIDENDINSKLINESIQNA 767
      : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |


```

```
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2961
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2961

Query Match      8.5%; Score 77; DB 4; Length 243;
Best Local Similarity 30.0%; Pred. No. 0.94;
Matches 36; Conservative 14; Mismatches 54; Indels 16; Gaps 5;

QY 58 GIGSFAQGDILGSLILGDAVGILITG-AYLDIKDFDNNNAKKADFKWTWKGMMLAG 116
Db 16 GINSFSLFILTAGWAAFPNCGIALLPSYISYL----IGETKDHSEFVAIFKGLGLGG 71
QY 117 VVTMAVTRLTETVLVPTFANNYRNKLNIAALGGPEPSFDINMQASA-LGFGLSFKK 175
Db 72 AMT-----TCFLTIFVLG-----LLIGGLGSALTGIFPILSLVWGIILALLGLGLMFGK 121

RESULT 5
US-08-840-466A-19
; Sequence 19, Application US/08840466A
; Patent No. 6261561
; GENERAL INFORMATION:
; APPLICANT: Stewart, C. Neal
;             McKee, Marian L.
;             O'Brien, Alison D.
;             Wachtel, Marian R.
; TITLE OF INVENTION: Method Of Stimulating An Immune Response
; By Administration Of Host Organisms That Express Intimin
; Alone Or As A Fusion Protein With One Or More Other
; Antigens.
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
;             Dunner, L.L.P.
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/840,466A
; FILING DATE: 18-Apr-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Boone, Laural S.
; REGISTRATION NUMBER: 43,505
; REFERENCE/DOCKET NUMBER: 04995.0029-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 934 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
```

```
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-08-840-466A-19

Query Match      8.2%; Score 74.5; DB 4; Length 934;
Best Local Similarity 21.7%; Pred. No. 6.7;
Matches 34; Conservative 22; Mismatches 46; Indels 55; Gaps 7;

QY 22 DSKSTENLGAGEFLVYETNKKDSLVPFLNLFLGFGIGSFAQGDILGSLILGDAVGI 81
Db 252 DSRETNALGAGQRF-----FLPANMLGYNV--FIDQFSGDNTRLG----- 290
QY 82 GLILTGAYLDIKDFDNNNAKKADFKW-TWKGMMLAGVVTMAVTRLTETVLVPTFANNYRN 140
Db 291 ---IGGEYW--RDYFKSSVNGYFRMRWHESY-----PALGAKLIYEQYY 354

QY 141 KLKNSLNIALGGPEPSFDINMQASALGFLSPKKS 177
Db 325 RPANGFDIRFNGYLPST-----PALGAKLIYEQYY 354

RESULT 6
US-09-696-188B-19
; Sequence 19, Application US/09696188B
; Patent No. 6406885
; GENERAL INFORMATION:
; APPLICANT: Stewart, C. Neal
;             McKee, Marian L.
;             O'Brien, Alison D.
;             Wachtel, Marian R.
; TITLE OF INVENTION: Method Of Stimulating An Immune Response
; By Administration Of Host Organisms That Express Intimi
; Alone Or As A Fusion Protein With One Or More Other
; Antigens.
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
;             Dunner, L.L.P.
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/696,188B
; FILING DATE: 26-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/840,466
; FILING DATE: 1997-04-18
; ATTORNEY/AGENT INFORMATION:
; NAME: Boone, Laural S.
; REGISTRATION NUMBER: 43,505
; REFERENCE/DOCKET NUMBER: 04995.0029-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 934 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-696-188B-19

Query Match      8.2%; Score 74.5; DB 4; Length 934;
Best Local Similarity 21.7%; Pred. No. 6.7;
```

| | | | | | | | | |
|--|-----|--|----------------|------------|-----|--------|-----|------|
| Matches | 34; | Conservative | 22; | Mismatches | 46; | Indels | 55; | Gaps |
| QY | 22 | DSKSTFNLAGAGEKFLVYTNNKKDSLVPFLNLFLGFGIGSFAQGDI | LGLGSLIFGDAYGI | 81 | | | | |
| Dd | 252 | DSRFETANILAGAQGRF-----FLPANMLGINV--FTDODFSGDNRLG----- | | 290 | | | | |
| QY | 82 | GLIITGAYLDTKDNDNNAKKADEKW-TWGKGMLGAVVTMAVTRLTETVLFFTANNYNR | 140 | | | | | |
| Dd | 291 | ---TGGEW--RDYFKSSVNGYFRMRWHESY-----HKKDYDE | 324 | | | | | |
| QY | 141 | KLKLSNLNTALGGFPSPEDINMQASALGFGLSFKKSY | 177 | | | | | |
| Dd | 325 | RPMANGEDIRFNGYLPSY-----PALCAKLIYEYY | 354 | | | | | |
| RESULT 7 | | | | | | | | |
| US-09-295-186-16 | | | | | | | | |
| ; Sequence 16, Application US/09295186B | | | | | | | | |
| ; Patent No. 6127137 | | | | | | | | |
| GENERAL INFORMATION: | | | | | | | | |
| ; APPLICANT: Hasida, Miyoko | | | | | | | | |
| ; APPLICANT: Tsutsumi, No. 6127137iko | | | | | | | | |
| ; APPLICANT: Halkier, Torben | | | | | | | | |
| ; APPLICANT: Stringer, Mary Ann | | | | | | | | |
| ; TITLE OF INVENTION: An Acidic Phospholipase, Production, and | | | | | | | | |
| ; TITLE OF INVENTION: Methods of Using Thereof (As Amended) | | | | | | | | |
| ; FILE REFERENCE: 4953.204-US | | | | | | | | |
| ; CURRENT APPLICATION NUMBER: US/09/295.186B | | | | | | | | |
| ; PRIORITY FILING DATE: 1999-04-20 | | | | | | | | |
| ; PRIOR APPLICATION NUMBER: 1215/96 | | | | | | | | |
| ; PRIOR FILING DATE: 1996-10-31 | | | | | | | | |
| ; PRIOR APPLICATION NUMBER: PCT/DK97/00490 | | | | | | | | |
| ; PRIOR FILING DATE: 1997-10-30 | | | | | | | | |
| ; NUMBER OF SEQ ID NOS: 18 | | | | | | | | |
| ; SOFTWARE: FastSeq for Windows Version 3.0 | | | | | | | | |
| ; SEQ ID NO 16 | | | | | | | | |
| ; LENGTH: 612 | | | | | | | | |
| ; TYPE: PRP | | | | | | | | |
| ; ORGANISM: Penicillium notatum | | | | | | | | |
| US-09-295-186-16 | | | | | | | | |
| Query Match 7.8%; Score 71; DB 3; Length 612; | | | | | | | | |
| Best Local Similarity 26.4%; Pred. No. 9.4; | | | | | | | | |
| Matches 33; Conservative 12; Mismatches 38; Indels 42; Gaps 5; | | | | | | | | |
| QY | 14 | AFSSFAODSKSTFNLAGAGEKFLVYTNNKKDSLVPFLNLFLGFGIGSFAQGDI----- | 67 | | | | | |
| Dd | 109 | AFDSDRTDNATATGHLLQLSATY-----ISGLSGGSWLGLSIINNFT | 152 | | | | | |
| QY | 68 | -----LGSSLILGEDAVGIGLI-LTGAYLDIKDFPDNNAKKADEFWT---- | 107 | | | | | |
| Dd | 153 | TVDKLPQTAEAGSVWFQGNIIIEGPDAggiQLLOSAGYKKDLADAVDGKKRAGFTTLTDI | 212 | | | | | |
| QY | 108 | WGKM 112 | | | | | | |
| Dd | 213 | WGRAL 217 | | | | | | |
| RESULT 8 | | | | | | | | |
| US-08-851-362D-37 | | | | | | | | |
| ; Sequence 37, Application US/08851362D | | | | | | | | |
| ; Patent No. 6235883 | | | | | | | | |
| GENERAL INFORMATION: | | | | | | | | |
| ; APPLICANT: Jakobovits, Aya | | | | | | | | |
| ; APPLICANT: Yang, Xiao-Dong | | | | | | | | |
| ; APPLICANT: Gallo, Michael | | | | | | | | |
| ; APPLICANT: Jia, Xiao-Chi | | | | | | | | |
| ; TITLE OF INVENTION: Human Monoclonal Antibodies to Epidermal | | | | | | | | |
| ; FILE REFERENCE: Cell 4.20 | | | | | | | | |
| ; CURRENT APPLICATION NUMBER: US/08/851,362D | | | | | | | | |
| ; PRIORITY FILING DATE: 1997-05-05 | | | | | | | | |
| ; NUMBER OF SEQ ID NOS: 49 | | | | | | | | |


```
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4175

Query Match          7.8%; Score 70.5; DB 4; Length 344;
Best Local Similarity 26.6%; Pred. No. 4.9;
Matches 33; Conservative 16; Mismatches 34; Indels 41; Gaps 5;

Qy 4 FLIFLIVIFCAFSPFAODSKSTFNL-----GAGEKFLVYETNKK-----D 44
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 209 YIIIVLVIV-----YANFNANAVENLLNIGSWIVWVGSSWISQIRLQAIKKQKRPD 263
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

Qy 45 SLVPP-----LLNLFGLFGIGSFAQCDILGSLILGLFSDAVGIGLILTGAY 89
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 264 KLLPKAPFPVPGVPIVLFLLLLGSSFGSIASGDILG--VIRNFTPLFLAIYIAH 321
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

Qy 90 LDIK 93
   ||
Db 322 KLIK 325
   ||

RESULT 11
US-09-393-627B-1
; Sequence 1, Application US/09393627B
; Patent No. 6455314
; GENERAL INFORMATION:
; APPLICANT: Wickham, Thomas J.
; APPLICANT: Kovesdi, Imre
; APPLICANT: Roelvinck, Petrus W.
; APPLICANT: Bruder, Joseph T.
; TITLE OF INVENTION: Alternatively Targeted Adenovirus
; FILE REFERENCE: 202345
; CURRENT APPLICATION NUMBER: US/09/393.627B
; CURRENT FILING DATE: 1999-09-10
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: US 60/099,851
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 60/136,529
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.2
; SEQ ID NO 1
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Human adenovirus serotype 5
US-09-393-627B-1

Query Match          7.8%; Score 70.5; DB 4; Length 581;
Best Local Similarity 26.2%; Pred. No. 10;
Matches 48; Conservative 18; Mismatches 72; Indels 45; Gaps 11;

Qy 22 DSKSTFNLGAGEKFLVYETNKKDSLVPFLNLFGLFGIGSFAQGDILG-----SII 73
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 215 DDLNLTLTATGPGVTINTNTSLQTKVG-----ALGFDQGNQNLNVAGGLRIDSONRRLI 269
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

Qy 74 LG-----FDA-----VGIG-LILTGAY-LDIK-----DFDNNAKKDFKWTWGKM 112
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 270 LDVSYFDAQNQLNRLGGQPLFINSAHNLDINYKGLYLFASNNKSKLEVNLSSTAKGL 329
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

Qy 113 ML-AGVVTMATRLTEIVLPFTFANNYRNKLNLSNIALG-GFEPSPFDINMQASALGFG 170
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 330 MFDATAINAGDGLGFGSP-----NAPNTNPLTKTIGHGLE--FDSNKAMVFKLTGTG 380
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

Qy 171 LSF 173
   |||
Db 381 LSF 383
   |||

RESULT 12
US-08-608-241-2
; Sequence 2, Application US/08608241
; Patent No. 5747328
; GENERAL INFORMATION:
; APPLICANT: Donohue, Timothy J
```

```
; APPLICANT: Barber, Robert D
; APPLICANT: Withuhn, Vernon
; TITLE OF INVENTION: MICROBIAL SYSTEM FOR FORMALDEHYDE
; TITLE OF INVENTION: SENSING AND REMEDIATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESS: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/608.241
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27,386
; REFERENCE/DOCKET NUMBER: 960296.93511
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-608-241-2

Query Match          7.6%; Score 69; DB 1; Length 376;
Best Local Similarity 25.3%; Pred. No. 8.2;
Matches 37; Conservative 23; Mismatches 54; Indels 32; Gaps 6;

Qy 48 PFLNLFGLFGIG-----SFAQGDILGSLILGLFSDAVGI-----GLILTGAYLDIKDPD 96
   ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 161 PFDKICYICGVTGTGAVINTAKVEIGAKAVFVGLGGIGLNVIOQLAGADMIIIGVDL 220
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

Qy 97 NNAKKADFQKWTGK-----GMLAGVVTMAVTRLEI-VLPFTFANNYRNK 141
   |||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 221 NNAKK-----EMGERFGMTHFVNPSEIDGVVVAHLVNMKTPTDFQIGGADYTFDCTGNVK 275
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

Qy 142 -LKNSLNIALGGFEPSPFDINMQASA 166
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 276 VMRQALEACHRGWGOSIVIGVAPAGA 301
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

RESULT 13
US-08-922-182-2
; Sequence 2, Application US/08922182
; Patent No. 5834300
; GENERAL INFORMATION:
; APPLICANT: Donohue, Timothy J
; APPLICANT: Barber, Robert D
; APPLICANT: Withuhn, Vernon
; TITLE OF INVENTION: MICROBIAL SYSTEM FOR FORMALDEHYDE
; TITLE OF INVENTION: SENSING AND REMEDIATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESS: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```


Search completed: March 14, 2003, 14:27:50
Job time : 13.9382 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:25:51 ; Search time 9.94382 Seconds
(without alignments)
820.439 Million cell updates/sec

Title: US-09-508-487-23
Perfect score: 908
Sequence: 1 MNKELIFILVIFCAFSFAQ.....DINMGASALGFLSFKKSY 177

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 88.5 | 9.7 | 451 | 10 | US-09-741-669-422 |
| 2 | 87 | 9.6 | 413 | 9 | US-09-910-186A-22 |
| 3 | 87 | 9.6 | 852 | 9 | US-10-011-588-25 |
| 4 | 81.5 | 9.0 | 206 | 10 | US-09-815-242-11462 |
| 5 | 78 | 8.6 | 544 | 9 | US-09-738-626-3927 |
| 6 | 78 | 8.6 | 544 | 10 | US-09-948-777-2 |
| 7 | 76.5 | 8.4 | 417 | 9 | US-09-738-626-6850 |
| 8 | 74.5 | 8.2 | 934 | 8 | US-08-837-459-19 |
| 9 | 73 | 8.0 | 406 | 9 | US-09-738-626-4686 |
| 10 | 73 | 8.0 | 421 | 9 | US-09-738-626-6813 |
| 11 | 72 | 7.9 | 670 | 9 | US-10-060-763-10 |
| 12 | 72 | 7.9 | 670 | 12 | US-10-063-763-10 |
| 13 | 71.5 | 7.9 | 513 | 9 | US-09-738-626-6860 |
| 14 | 70.5 | 7.8 | 95 | 9 | US-09-187-693-53 |
| 15 | 70.5 | 7.8 | 410 | 10 | US-09-767-041-20 |
| 16 | 70.5 | 7.8 | 579 | 10 | US-09-870-203A-6 |
| 17 | 70.5 | 7.8 | 579 | 10 | US-09-870-203A-8 |
| 18 | 70.5 | 7.8 | 579 | 10 | US-09-870-203A-18 |
| 19 | 70.5 | 7.8 | 579 | 10 | US-09-870-203A-20 |

| | | | | | |
|----|------|-----|------|----|---------------------|
| 20 | 70.5 | 7.8 | 580 | 10 | US-09-870-203A-10 |
| 21 | 70.5 | 7.8 | 580 | 10 | US-09-870-203A-12 |
| 22 | 70.5 | 7.8 | 581 | 10 | US-09-758-008-5 |
| 23 | 70.5 | 7.8 | 581 | 10 | US-09-870-203A-2 |
| 24 | 70.5 | 7.8 | 581 | 10 | US-09-870-203A-4 |
| 25 | 70.5 | 7.8 | 581 | 10 | US-09-870-203A-14 |
| 26 | 70.5 | 7.8 | 581 | 10 | US-09-870-203A-16 |
| 27 | 70.5 | 7.8 | 600 | 10 | US-09-871-212-5 |
| 28 | 70.5 | 7.8 | 1161 | 8 | US-08-910-386A-20 |
| 29 | 70 | 7.7 | 452 | 10 | US-09-815-242-10183 |
| 30 | 69.5 | 7.7 | 650 | 10 | US-09-815-242-13341 |
| 31 | 68.5 | 7.5 | 154 | 10 | US-09-881-752A-118 |
| 32 | 68.5 | 7.5 | 475 | 9 | US-09-738-626-6117 |
| 33 | 68 | 7.5 | 255 | 10 | US-09-881-752A-106 |
| 34 | 68 | 7.5 | 276 | 9 | US-09-895-913A-356 |
| 35 | 68 | 7.5 | 570 | 9 | US-09-344-882-28 |
| 36 | 68 | 7.5 | 570 | 12 | US-10-062-254-367 |
| 37 | 68 | 7.5 | 691 | 10 | US-09-841-132-313 |
| 38 | 68 | 7.5 | 2039 | 9 | US-10-192-584-7 |
| 39 | 68 | 7.5 | 2042 | 9 | US-10-192-584-6 |
| 40 | 67 | 7.4 | 132 | 10 | US-09-800-729-134 |
| 41 | 67 | 7.4 | 132 | 10 | US-09-800-729-189 |
| 42 | 67 | 7.4 | 156 | 10 | US-09-800-729-135 |
| 43 | 67 | 7.4 | 209 | 9 | US-09-813-453A-21 |
| 44 | 67 | 7.4 | 239 | 10 | US-09-840-787-22 |
| 45 | 67 | 7.4 | 239 | 10 | US-09-800-729-102 |

ALIGNMENTS

RESULT 1

US-09-741-669-422
; Sequence 422, Application US/09741669
; Patent No. US20020022718A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; TITLE OF INVENTION: Genes identified as required for proliferation of E. coli
; FILE REFERENCE: ELITRA.009A
; CURRENT APPLICATION NUMBER: US/09/741.669
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 422
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-741-669-422

| | | | | |
|-----------------------|-------|--|--------|----------------|
| Query Match | 9.7% | Score 88.5; | DB 10; | Length 451; |
| Best Local Similarity | 27.5% | Pred. No. 0.18; | | |
| Matches | 30; | Conservative | 15; | Mismatches 33; |
| | | | | Indels 31; |
| | | | | Gaps 6; |
| QY | 46 | LVPFLLNLF---LGFIGSFAQDILGSLILGFDVAGVIGILTGAYLD-----IK 93 | | |
| Db | 17 | MLPIVIIFSKILG-----MKAGDCKRAGLHIGIFGVIGLVI-GLMLDSIGPAKAAAE 70 | | |
| QY | 94 | DFDNNAKADPKW-----TWCKGMMLAGV-----VTMAVTRLTEIV 129 | | |
| Db | 71 | NFDNLNHHVDVYWGPGSPMTWASQIALVAIPAILNVNVAMLLTRRVV 119 | | |

RESULT 2

US-09-910-186A-22
; Sequence 22, Application US/09910186A
; Publication No. US20030009025A1
; GENERAL INFORMATION:
; APPLICANT: U.S. Army Medical Research & Material Command

```
; TITLE OF INVENTION: RECOMBINANT VACCINE AGAINST BOTULINUM
; FILE OF INVENTION: NEUROTOXIN
; FILE REFERENCE: A33626-A 067252.0107
; CURRENT APPLICATION NUMBER: US/09/910,186A
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: PCT/US00/12890
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: 09/611,419
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 60/133,865
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/133,866
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/133,867
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/133,868
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/133,869
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/133,873
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 08/123,975
; PRIOR FILING DATE: 1993-09-21
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-09-910-186A-22
```

```
Query Match          9.6%; Score 87; DB 9; Length 413;
Best Local Similarity 22.8%; Pred. No. 0.24;
Matches 38; Conservative 31; Mismatches 58; Indels 40; Gaps 7;

QY 37 VYETNKKDS-----LVPEL---LNFLGFGIGSFAQG-DILGGSILILGFD-----AV 79
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 162 VIEANKSNTMDKIADISLIVPYIGLALNVGNETAKGNFENAFIAGASILLEFIPELLIP 221
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 80 GIGLIITGAYLD-----IKDFDNNAKKADFKWTGKGMMLAGVVTMAVTRL----- 125
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 222 VVGAFLESYIDNKKIITIDNALTKRNEKWSMDYGLIVAQWLSTVNTQFYTIKEGMKY 281
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 126 -----TEIVLPETFANNYRKLNLSNIALGGFEPSPFDINMGOA 164
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 282 ALNYQAALIEEIKYRY-NIYSEKESNINIDFNDSKLNKLNINQ 327
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
```

```
RESULT 3
US-10-011-588-25
; Sequence 25, Application US/10011588
; Patent No. US20020168727A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Leonard
; TITLE OF INVENTION: RECOMBINANT LIGHT CHAINS OF BOTULINUM
; TITLE OF INVENTION: NEUROTOXINS AND LIGHT CHAIN FUSION PROTEINS FOR USE IN
; TITLE OF INVENTION: RESEARCH AND CLINICAL THERAPY
; FILE REFERENCE: A34796 067252.0113
; CURRENT APPLICATION NUMBER: US/10/011,588
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 09/910,186
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 09/611,419
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 60/246,744
; PRIOR FILING DATE: 2000-11-06
; PRIOR APPLICATION NUMBER: 60/311,966
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 25
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant protein encoded by SEQ ID NO:24
US-10-011-588-25

Query Match          9.6%; Score 87; DB 9; Length 852;
Best Local Similarity 22.8%; Pred. No. 0.56;
Matches 38; Conservative 31; Mismatches 58; Indels 40; Gaps 7;

QY 37 VYETNKKDS-----LVPEL---LNFLGFGIGSFAQG-DILGGSILILGFD-----AV 79
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 601 VIEANKSNTMDKIADISLIVPYIGLALNVGNETAKGNFENAFIAGASILLEFIPELLIP 660
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 80 GIGLIITGAYLD-----IKDFDNNAKKADFKWTGKGMMLAGVVTMAVTRL----- 125
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 661 VVGAFLESYIDNKKIITIDNALTKRNEKWSMDYGLIVAQWLSTVNTQFYTIKEGMKY 720
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 126 -----TEIVLPETFANNYRKLNLSNIALGGFEPSPFDINMGOA 164
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 721 ALNYQAALIEEIKYRY-NIYSEKESNINIDFNDSKLNKLNINQ 766
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 4
US-09-815-242-11462
; Sequence 11462, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11462
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11462

Query Match          9.0%; Score 81.5; DB 10; Length 206;
Best Local Similarity 23.6%; Pred. No. 0.38;
Matches 33; Conservative 21; Mismatches 37; Indels 49; Gaps 6;

QY 60 GSAFGDILGGSILILGFDVAGIGLIITGAYLDIKDFDNNAKKADFKWTGKGMMLAGV-- 117
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 86 GHFVQGI-----DAIGV-----TEKIHANQVDFISASETLLLCVEQ 126
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
```


QY 121 AVTRLTEIVLPFTFANNYNRKLNLSNIALGGFEP-----S 156
| : | | | : | | : | | : | |
Db 248 LGTSSAATIPVT---YQTLKNDVDVNVAGFVPLCATIHLGSMKIGLFTFAVVM 303
| : | | : | | : | | : | |
QY 157 FDNMQASALGF 169
| : | | : | | : | | : | |
Db 304 YDMEVGVLISIGF 316
| : | | : | | : | | : | |
RESULT 8
US-08-837-459-19
; Sequence 19, Application US/08837459
; Publication No. US20020006407A1
; GENERAL INFORMATION:
; APPLICANT: McKee, Marian L.
; APPLICANT: O'Brien, Alison D.
; APPLICANT: Wachtel, Marian R.
; TITLE OF INVENTION: Histidine-Tagged Intimin And Methods Of
; USING Intimin To Stimulate An Immune Response And As An
; TITLE OF INVENTION: Antigen Carrier With Targeting Capability.
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner, L.B.P.
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/837,459
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32,984
; REFERENCE/DOCKET NUMBER: 04995-0023-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 934 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-837-459-19
Query Match 8.2%; Score 74.5; DB 8; Length 934;
Best Local Similarity 21.7%; Pred. No. 13;
Matches 34; Conservative 22; Mismatches 46; Indels 55; Gaps 7;
QY 22 DSKSTNLGAGEKFLVYETNKKDSLPVFLNLFLGFGISFAGDILGSLILGFDVAGI 81
| : | | | | : | | : | | : | |
Db 252 DSRFTANLGAGORF-----FLPANMLGYNV--FIDQDFSGDNRILG----- 290
| : | | | | : | | : | | : | |
QY 82 GLILTAYLDIDFNNNAKKADFK-TWKGMMLAGVVTMAVTRLTEIVLPFTFANNYNR 140
| : | | : | | : | | : | | : | |
Db 291 ---IGGEYV--RDYKSSVNGVFRMRWHESY-----HKDYDE 324
| : | | : | | : | | : | | : | |
QY 141 KLNLSNIALGGFEPFSDINMGQASALGFLGFKKSY 177
| : | | : | | : | | : | | : | |
Db 325 RPANGFDIRENGVLPSTY-----PALGAKLIVEQY 354
| : | | : | | : | | : | | : | |
RESULT 9

US-09-738-626-4686
; Sequence 4686, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent In ver. 3.0
; SEQ ID NO 4686
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4686
Query Match 8.0%; Score 73; DB 9; Length 406;
Best Local Similarity 24.6%; Pred. No. 6.6;
Matches 46; Conservative 26; Mismatches 53; Indels 62; Gaps 10;
QY 6 IFILVIFCAPS-----SFAQDDSKSTNLGAGEKFLVYETNKKDSLPVFLNLFLG-- 56
| | | | : | | : | | : | |
Db 218 IFILLCFAYSGVIAYINAFABERDLIT---GAGLFFIAYAVSM-----FVNRSLGKL 268
| | | | : | | : | | : | |
QY 57 -----FGIGSFAQDDILGSLILGFDVAGIILTGAYLDIKDFDNNKADFK 105
| | | | : | | : | | : | |
Db 269 QDRRGDNVVIYFGLFF-----VISITLSFATSNHVVLSGVIAGL----- 310
| | | | : | | : | | : | |
QY 106 WTWKGMMLAGVVTMAVTRLTEIVLPFTFANNYNRKLNLSNIALG-GFEPSPDINMGQA 164
| : | | : | | : | | : | | : | |
Db 311 ---GYGLMPAVQSIAGVGVDTKTECTATSTLF-----LFVDLGFGRGP---IILGAV 357
| : | | : | | : | | : | | : | |
QY 165 S-ALGFG 170
| : | | : | | : | | : | |
Db 358 SAAIGFG 364
| : | | : | | : | | : | |
RESULT 10
US-09-738-626-6813
; Sequence 6813, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:21:36 ; Search time 7.33333 Seconds
(without alignments)
1006.744 Million cell updates/sec

Title: US-09-508-487-21

Perfect score: 904

Sequence: 1 MNKFLIVLLAFVCFSSFAQ.....DINMQASALGFELSPKSKY 178

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|-------------|
| 1 | 86.5 | 9.6 | 545 | 1 | NU5M_ALBCO |
| 2 | 85 | 9.4 | 538 | 1 | TCMA_STRGA |
| 3 | 83.5 | 9.2 | 442 | 1 | DTA2_RHILO |
| 4 | 83.5 | 9.2 | 449 | 1 | YKE6_YEAST |
| 5 | 81.5 | 9.0 | 692 | 1 | PHSG_AQUAE |
| 6 | 81.5 | 9.0 | 1290 | 1 | BXB_CLOBO |
| 7 | 81 | 9.0 | 507 | 1 | YQGP_BACSU |
| 8 | 80.5 | 8.9 | 371 | 1 | COXB_AGR4 |
| 9 | 80 | 8.8 | 750 | 1 | CBBB_BACTV |
| 10 | 79.5 | 8.8 | 419 | 1 | CMLA_PSEAE |
| 11 | 78.5 | 8.7 | 436 | 1 | DTA_PSEAE |
| 12 | 78.5 | 8.7 | 663 | 1 | CYOB_ECOLI |
| 13 | 78 | 8.6 | 235 | 1 | CCDA_BACSU |
| 14 | 77.5 | 8.6 | 444 | 1 | DTA_RHILO |
| 15 | 76.5 | 8.5 | 933 | 1 | SLAP_CAMFE |
| 16 | 74.5 | 8.2 | 264 | 1 | YA96_MYCPN |
| 17 | 74.5 | 8.2 | 429 | 1 | DTA_YERPE |
| 18 | 74 | 8.2 | 578 | 1 | AC22_STRCO |
| 19 | 74 | 8.2 | 1113 | 1 | IMB4_YEAST |
| 20 | 73.5 | 8.1 | 346 | 1 | TERC_ALCSP |
| 21 | 73.5 | 8.1 | 490 | 1 | MODF_ECOLI |
| 22 | 73.5 | 8.1 | 992 | 1 | A1A4_HUMAN |
| 23 | 73.5 | 8.1 | 1032 | 1 | A1A4_MOUSE |
| 24 | 73 | 8.1 | 263 | 1 | LPXA_CAMJE |
| 25 | 73 | 8.1 | 473 | 1 | MOT4_CHICK |
| 26 | 72.5 | 8.0 | 230 | 1 | CRP_RAT |
| 27 | 72.5 | 8.0 | 451 | 1 | PTKC_ECOLI |
| 28 | 72.5 | 8.0 | 470 | 1 | PTTB_BACSU |
| 29 | 72.5 | 8.0 | 516 | 1 | YWCA_BACSU |
| 30 | 72.5 | 8.0 | 631 | 1 | PTBA_ERWCH |
| 31 | 72 | 8.0 | 393 | 1 | NCAP_PVM |
| 32 | 72 | 8.0 | 477 | 1 | YPUM_RHOCA |
| 33 | 72 | 8.0 | 634 | 1 | ICFG_SYNY3 |

34 71.5 7.9 151 1 YQ00_MYCPA Q9K537 mycobacteri
35 71.5 7.9 489 1 Y225_MYCGE P47467 mycoplasma
36 71.5 7.9 620 1 KEFC_ECOLI P03819 escherichia
37 71.5 7.9 809 1 NAH2_RABIT P50482 oryctolagus
38 71.5 7.9 812 1 NAH2_HUMAN Q9UBY0 homo sapien
39 71.5 7.9 845 1 CC47_YEAST P38132 saccharomyc
40 71.5 7.9 922 1 PMP1_CHLPN Q9Z9G5 chlamydia p
41 71.5 7.9 1077 1 HGPA_HAEN Q9ZAZ1 haemophilus
42 71 7.9 172 1 RL10_RHILO Q98N68 rhizobium l
43 71 7.9 406 1 RENI_HUMAN P00797 homo sapien
44 71 7.9 649 1 COX1_BACSU P34956 bacillus su
45 71 7.9 660 1 CCMF_BRAJA P45403 bradyrhizob

ALIGNMENTS

RESULT 1
NU5M_ALBCO
ID NU5M_ALBCO STANDARD; PRT; 545 AA.
AC P48918:
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 01-FEB-1996 (Rel. 33, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3).
GN ND5.
OS Albinaria coerulea (Land snail).
OC Mitochondrion.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC Clausiliidae; Alopinae; Albinaria.
OX NCBI_TaxID=42349;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96120351; PubMed=7498775;
RA Hatzoglou E., Rodakis G.C., Lekanidou R.;
RT "Complete sequence and gene organization of the mitochondrial genome
of the land snail Albinaria coerulea";
RL Genetics 140:1353-1366(1995).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed, usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X83390; CAA58296.1; -
DR InterPro; IPR003916; NADHub_Oxred5.
DR InterPro; IPR001750; Oxidored_q1.
DR InterPro; IPR001516; Oxidored_q1_N.
DR Pfam; PF00361; oxidored_q1; 1.
DR Pfam; PF00662; oxidored_q1_N; 1.
DR PRINTS; PR01434; NADHDHGNASE5.
DR KX Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 545 AA; 61335 MW; BE801FDEC7A18563 CRC64;

Query Match 9.6%; Score 86.5; DB 1; Length 545;
Best Local Similarity 29.1%; Pred. No. 2;
Matches 39; Conservative 21; Mismatches 37; Indels 37; Gaps 7;
QY 4 FLIVVLL-AFCVFSFQAADSKSAFNLCAGEKLLAYETSKKDPVPLNLLFGFGIGS 62
Db 51 FLVVLVLISSCVF-LFANEYMSDHYNIRFGWLIISFVAS-----MGILLIS 96
QY 63 FAQGDILGGFLITGLFDVAGVIGLITGLAYDILKALDKAPKAFAKWTGCKMMLAGAVTMA 122
Db 97 ---GSIF--TLLGWDGLGITSFALIAYD-----NYNASSAF-----LFLM 134
QY 123 VTRLTEIIPFTFA 136
Db 135 TNRIGDVLIIATFS 148

```

RESULT 2
TCMA_STRGA          STANDARD;          PRT;          538 AA.
AC P39886;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tetracenomycin C resistance and export protein.
GN TCMA.
OS Streptomyces glaucescens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinobacteriales; Streptomycetales; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1907;
RN [1]
RP STRAIN=ATCC 13032 / ETH 22794 / GLA.0;
RC MEDLINE=92276347; PubMed=1592819;
TX Guilfoile P.G., Hutchinson C.R.;
RT "Sequence and transcriptional analysis of the Streptomyces glaucescens tcmAR tetracenomycin C resistance and repressor gene loci."
RT J. Bacteriol. 174:3651-3658(1992).
CC -!- FUNCTION: RESISTANCE TO TETRACENOMYCIN C BY AN ACTIVE TETRACENOMYCIN C EFFLUX SYSTEM WHICH IS PROBABLY ENERGIZED BY TRANSMEMBRANE ELECTROCHEMICAL GRADIENTS.
CC -!- PATHWAY: Polyketide antibiotic tetracenomycin C biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSLUCASE FAMILY).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
CC EMBL; M80674; AAA67509.1;
CC PIR; S27687; S27687.
CC PIR; A41901; A41901.
CC InterPro; IPR004638; Efflux_EmrB.
CC InterPro; IPR003662; sub transporter.
CC Pfam; PF00083; sugar_tr; 1.
CC TIGRFAMs; TIGR00711; efflux_EmrB; 1.
CC Antibiotic resistance; Antibiotic biosynthesis; Transmembrane; Transport.
CC TRANSMEM 28 48 POTENTIAL.
CC TRANSMEM 65 85 POTENTIAL.
CC TRANSMEM 100 120 POTENTIAL.
CC TRANSMEM 126 146 POTENTIAL.
CC TRANSMEM 154 174 POTENTIAL.
CC TRANSMEM 181 201 POTENTIAL.
CC TRANSMEM 213 233 POTENTIAL.
CC TRANSMEM 239 259 POTENTIAL.
CC TRANSMEM 286 306 POTENTIAL.
CC TRANSMEM 319 339 POTENTIAL.
CC TRANSMEM 342 362 POTENTIAL.
CC TRANSMEM 371 391 POTENTIAL.
CC TRANSMEM 413 433 POTENTIAL.
CC TRANSMEM 494 514 POTENTIAL.
CC SEQUENCE 538 AA; 54846 MW; DED3F28C1F22AA56 CRC64;
Query Match
Best Local Similarity 9.4%; Score 85; DB 1; Length 538;
Matches 33; Conservative 19; Mismatches 38; Indels 44; Gaps 6;
QY 49 PFLLNLFLGFGIG----SFAQGDILGGFLI--LGFDA-----VCIGILITGAYLIDIKA 95
DB 145 PKLNLWAIIGWSGVGASTAGPIIGGLLVQHVGHNEAFINVPGLAALVAGLVILTDA 204

```


AC P10844; P10843;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type B precursor (EC 3.4.24.69) (BoNT/B)
 DE (bontoxilysin B).
 GN BOTB.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92384550; PubMed=1514783;
 RA Whelan S.M., Elmore M.J., Bodsworth N.J., Brehm J.K., Atkinson T.,
 RA Minton N.P.;
 RT "Molecular cloning of the Clostridium botulinum structural gene
 RT encoding the type B neurotoxin and determination of its entire
 RT nucleotide sequence.";
 RT Appl. Environ. Microbiol. 58:2345-2354(1992).
 [2]
 RP SEQUENCE OF 35-245 FROM N.A.
 RC STRAIN=NCTC 7273;
 RA Szabo E.A., Pemberton J.M., Desmarchelier P.M.;
 RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE OF 633-993 FROM N.A.
 RC STRAIN=NCTC 7273;
 RX MEDLINE=94013372; PubMed=8408542;
 RA Campbell K., East A.K., Collins M.D.;
 RT "Gene probes for identification of the botulin neurotoxin gene and
 RT specific identification of neurotoxin types B, E, and F.";
 RL J. Clin. Microbiol. 31:2255-2262(1993).
 [4]
 RP SEQUENCE OF 1-44 AND 441-466.
 RC STRAIN=657;
 RX MEDLINE=69000987; PubMed=3139097;
 RA Dasgupta B.R., Datta A.;
 RT "Botulinum neurotoxin type B (strain 657): partial sequence and
 RT similarity with tetanus toxin.";
 RL Biochimie 70:811-817(1988).
 [5]
 RP SEQUENCE OF 1-16 AND 441-458.
 RC STRAIN=OKRA;
 RX MEDLINE=85197963; PubMed=3888113;
 RA Schmidt J.J., Sathyamoorthy V., Dasgupta B.R.;
 RT "Partial amino acid sequences of botulinum neurotoxins types B and
 RT E.";
 RL Arch. Biochem. Biophys. 238:544-548(1985).
 [6]
 RP IDENTIFICATION AS ZINC-PROTEASE.
 RX MEDLINE=93054694; PubMed=1429690;
 RA Schiavo G., Rossetto O., Santucci A., Dasgupta B.R., Montecucco C.;
 RT "Botulinum neurotoxins are zinc proteins.";
 RL J. Biol. Chem. 267:23479-23483(1992).
 [7]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=93063293; PubMed=1331807;
 RA Schiavo G., Benfenati F., Poulin B., Rossetto O., de Laureto P.P.,
 RA Dasgupta B.R., Montecucco C.;
 RT "Tetanus and botulinum B neurotoxins block neurotransmitter release
 RT by proteolytic cleavage of synaptobrevin.";
 RL Nature 359:832-835(1992).
 CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CLEAVES THE 76-GLN-|-PHE-77 BOND OF
 CC SYNAPTOSOMAL-2.
 CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.

CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M81186; AAA23211.1; -;
 CC EMBL; Z11934; CAA77991.1; -;
 CC EMBL; X70817; CAA50148.1; -;
 CC PIR; S07128; S07128.
 CC PIR; S07155; S07155.
 CC PIR; S08562; S08562.
 CC PIR; S08573; S08573.
 CC PIR; S08574; S08574.
 CC PIR; A48940; A48940.
 CC HSSP; P10845; 3BTA.
 CC MEROPS; M27_002; -;
 CC InterPro; IPR000395; Bontoxilysin.
 CC InterPro; IPR000130; Zn_MTPeptidse.
 CC Pfam; PF01742; Peptidase_M27; 1.
 CC PRINTS; PR00760; BONTOXILYSIN.
 CC ProDom; PD001963; Bontoxilysin; 1.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 KW INIT_MET 0
 FT CHAIN 1 440 BOTULINUM NEUROTOXIN B, LIGHT-CHAIN.
 FT METAL 441 1290 BOTULINUM NEUROTOXIN B, HEAVY-CHAIN.
 FT ACT_SITE 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 230 230 BY SIMILARITY.
 FT METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 436 445 INTERCHAIN (PROBABLE).
 FT CONFLICT 29 29 T -> M (IN REF. 4).
 FT CONFLICT 217 217 R -> G (IN REF. 2).
 FT CONFLICT 224 224 A -> S (IN REF. 2).
 FT CONFLICT 463 463 S -> R (IN REF. 4).
 SQ SEQUENCE 1290 AA; 150670 MW; D21746E2C024DF43 CRC64;
 Query Match 9.0%; Score 81.5; DB 1; Length 1290;
 Best Local Similarity 23.8%; Pred. No. 13;
 Matches 36; Conservative 24; Mismatches 56; Indels 35; Gaps 6;
 QY 47 IVPFLNLFGLFGIGSFAOGDILGFLILGFDV-----GIGLILTGAYLD--- 92
 Db 619 IVPY-IGLALNVG-NETAGNFENAEIAGASILLFEPILLIPVVGAFLESYIDNKK 676
 QY 93 -IKALDNAPKAAFKWGWKMLAGAVTMAVTRL-----TEIIPFT 134
 Db 677 IIKTDNALTKRNEKWSMDYGLIVAGLWLSVTNFTQFYTIKGYKALNYQAQALEIKYR 736
 QY 135 FANSYWRKLNKSLNIAFGGFEPSFDINMQOA 165
 Db 737 Y-NIYSEKESKNIDFNDINSKLNENGINQA 766
 RESULT 7
 YQGP_BACSU
 ID YQGP_BACSU STANDARD; PRT; 507 AA.
 AC P54493;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein yqgp.

| | |
|--|---|
| | 9.0%; Score 81; DB 1; Length 507; |
| Query Match | |
| Best Local Similarity | 28.3%; Pred.No. 5.8; |
| Matches | 30; Conservative 17; Mismatches 37; Indels 22; Gaps 5; |
| <hr/> | |
| QY 3 KFLIVL-----LAFCVSFSSFAQAADSSAFN-LGA-----GEKLLAYETSKKDPIV 48 | : :: : : |
| Dd 260 RFLLIYAAGITGSIAFSVPSPSAGASCAIGCGLCGALLYVALSNRKMFLRTGTNTIIV 319 | : :: : : |
| <hr/> | |
| QY 49 PFLNLPLFGFSGFAQ-----GDILGGFLILGFDAVGIGLILTGay 90 | ::: :: : : |
| Dd 320 IILIINLGFGAVNSIDNSGHIGGLIGGF----FAAAALGPLKGAF 361 | : |
| <hr/> | |
| RESULT 8 | |
| OXXB_AGRT4 | STANDARD; PRt; 371 AA. |
| ID OOXB_AGRt4 | |
| AC Q59159; | |
| Dt DT | 30-MAY-2000 (Rel. 39, Created) |
| Dt DT | 30-MAY-2000 (Rel. 39, Last sequence update) |
| Dt Dt | 15-JUN-2002 (Rel. 41, Last annotation update) |
| DE DE | Opine oxidase subunit B (EC 1.-.-) (Octopine oxidase subunit B). |
| Gn GN | Oxob. |
| OS Agrobacterium tumefaciens (strain Ach5). | |
| Og Og | Plasmid pTiAch5, |
| Oc Oc | Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group; |
| Oc OC | Rhizobiaceae; Rhizobium. |
| OX NCBI_TaxId=176298; | |
| RN RN | [1] |
| RP RP | SEQUENCE FROM N.A. |
| RA RA | MEDLINE=94321320; PubMed=8045881; |
| RR RR | Zanker H., Lurz G., Langridge U., Langridge P., Kreusch D., |
| Schroeder J.; | "Octopine and nopaline oxidases from Ti plasmids of Agrobacterium |
| tumefaciens: molecular analysis, relationship, and functional | characterization."; RT RT |
| J. Bacteriol. 176:4511-4517(1994). | RL RL |
| - - FUNCTION: OXIDATIVE CLEAVAGE OF OCTOPINE INTO L-ARGININE AND | CC CC |
| PYRUvATE (BY SIMILARITY). | CC CC |
| - - PATHWAY: CATABOLIC UTILIZATION OF OCTOPINE. | CC CC |
| - - SUBUNIT: HETERODIMER OF A SUBUNIT AND A B SUBUNIT. | CC CC |
| This SWISS-PROT entry is copyright. It is produced through a collaboration | CC CC |
| between the Swiss Institute of Bioinformatics and the EMBL Outstation - | CC CC |
| the European Bioinformatics Institute. There are no restrictions on its | CC CC |
| use by non-profit institutions as long as its content is in no way | CC CC |
| modified and this statement is not removed. Usage by and for commercial | CC CC |
| entities requires a license agreement (see http://www.isb-sib.ch/announce/ | CC CC |
| or send an email to licensel@isb-sib.ch). ----- | CC CC |
| EMBL; Z30328; CAAB2987.1; .. | DR DR |
| InterPro; IPR000927; Daa_Oxidase. | DR DR |
| InterPro; IPR000205; NAD_binding. | DR DR |
| Pfam; PF01266; DAO; 1. | KW KW |
| XOxidoreductase; Plasmid. | |
| SEQUENCE 371 AA; 39939 MW; 948705136822AC52 CRC64; | SQ SQ |
| <hr/> | |
| Query Match | 8.9%; Score 80.5; DB 1; Length 371; |
| Best Local Similarity | 30.9%; Pred.No. 4.7; |
| Matches | 38; Conservative 13; Mismatches 39; Indels 33; Gaps 6; |
| <hr/> | |
| QY 68 ILGGFLILGFDAVGIG-----LILTGAYLDIKALDNAPKAFTWTWGKMML----- 115 | :: :: : : |
| Dd 8 II GGGLYGASTAWLARSGTKPLYLDGADLDLRAN---FALVVQGKGLHAPHYALW 64 | :: :: : : |
| <hr/> | |
| QY 116 --AGA--VTMAYTRLTEII-----PTFEANS-----YNKLKNLSLIAFGGFPES 157 | :: :: : : |
| Dd 65 SDASARRWPMTANTLLDDSDGVLDQQDGAFTFALSEELEANRMDSIELTNGRAPQ 124 | :: :: : : |
| Oy Oy | 158 FDI 160 |

Db 125 FEB 127

RESULT 9

CBBB_BACTV STANDARD; PRT; 750 AA.
AC Q92IU5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Pesticidal crystal protein cryIIbB (insecticidal delta-endotoxin
DE CRYIIB(b)) (Crystalline entomocidal protoxin) (84 kDa crystal protein).
GN CRYIIBB OR CRYIIB(b).
OS Bacillus thuringiensis (subsp. medellin).
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
OX NCBI_TaxID=79672;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CIB 163-131;
MEDLINE=98449784; PubMed=9774743;
RA Orduz S., Realpe M., Arango R., Murillo L.A., Delecluse A.;
RT "Sequence of the cryIIbB1 gene from Bacillus thuringiensis subsp.
RT medellin and toxicity analysis of its encoded protein.";
RL Biochim. Biophys. Acta 1388:267-272(1998).
CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF MOSQUITOS SUCH AS. Aedes Aegypti, ANOPHELES
CC ALBIMANUS AND CULEX QUINQUEFASCIATUS LARVAE.
CC -!- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
CC SPOROGATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
CC OF THE SPORE COAT.
CC -!- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE
CC N-TERMINUS.
CC -!- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF017416; AAC97162.1; -;
CC InterPro; IPR001178; Endotoxin.
CC Pfam; PF00555; endotoxin; 1.
CC Toxin; Sporulation.
CC SEQUENCE 750 AA; 84407 MW; 293C383018B33D46 CRC64;
CC -----

Query Match 8.8%; Score 80; DB 1; Length 750;
Best Local Similarity 20.3%; Pred. No. 10;
Matches 37; Conservative 34; Mismatches 83; Indels 28; Gaps 6;
QY 15 FSSFAQADSKSFAFNLCAGEKLLAYTSKK-DPIVFLNLFGLFGIGFPAQGD----IL 69
Db 523 FTTVLGNKAKYSIRLTG-----FNTATRYRLVIRFKATARLAAGIRVSQSGNNRL 577
QY 70 GGFLLILG-----FDVAVGILGLITGAYLDIKALDKNAPKAFKWTGKGMMLA 116
Db 578 GGIPVEGNSGWDYIYDSTFNDLGITASTNAFFSI---DSDGVNASQWYLSKILLYK 634
QY 117 GAVTMVATRLTEIIFFTFANSYNRKLNLSNTAFGFFPFSDINMGQASALGFSLFKK 176
Db 635 DFVNNSGFRNQVPLAPVIAKPCNTFFVS--NNTSSGYEQGYNDNTQNTSSGYEQGYND 692
QY 177 SY 178
Db 693 NY 694

RESULT 10

CMLA_PSEAF

ID CMLA_PSEAF STANDARD; PRT; 419 AA.

AC P32482; Q56411; Q56412;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chloramphenicol resistance protein.
GN CMLA.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RX TRANSPOSON-Tn1696;
MEDLINE=91294195; PubMed=1648560;
RA Bissonnette L., Champetier S., Buissson J.-P., Roy P.H.;
RT "Characterization of the nonenzymatic chloramphenicol resistance
RT (cmla) gene of the In4 integron of Tn1696: similarity of the product
RT to transmembrane transport proteins.";
RL J. Bacteriol. 173:4493-4502(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX TRANSPOSON-Tn1696;
MEDLINE=92052679; PubMed=1658833;
RA Stokes H.W., Hall R.M.;
RT "Sequence analysis of the inducible chloramphenicol resistance
RT determinant in the Tn1696 integron suggests regulation by
RT translational attenuation.";
RL Plasmid 26:10-19(1991).
RN [3]
RP REVISIONS.
RC TRANSPOSON-Tn1696;
RA Partridge S., Hall R.M.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: IT APPEARS TO PROVOKE A REDUCTION OF THE CONTENT OF
CC THE MAJOR PORINS OMPA AND OMPC.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -!- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
CC AS THE DRUG RESISTANCE TRANSLOCASE FAMILY). BELONGS TO THE
CC BCR/CMLA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U12338; AAB60004.1; -;
CC EMBL; M64556; AAA26057.1; -;
CC PIR; A47033; A47033.
CC InterPro; IPR004734; Drug_resist.
CC InterPro; IPR004812; Efflux_Bcr_Cfla.
CC Pfam; PF003662; sub_transporter.
CC Pfam; PF00083; sugar_tr; 1.
CC TIGRFAMS; TIGR00710; efflux_Bcr_Cfla; 1.
CC TIGRFAMS; TIGR00880; 2_A_01_02_1.
KW Antibiotic resistance; Transport; Transmembrane; Inner membrane;
KW Plasmid; Transposable element.
FT DOMAIN 1 9 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 10 30 1 (POTENTIAL).
FT DOMAIN 31 46 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 47 67 2 (POTENTIAL).
FT DOMAIN 68 80 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 81 101 3 (POTENTIAL).
FT DOMAIN 102 109 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 110 127 4 (POTENTIAL).
FT DOMAIN 128 141 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 142 162 5 (POTENTIAL).
FT DOMAIN 163 169 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 170 187 6 (POTENTIAL).
FT DOMAIN 188 214 EXTRACELLULAR (POTENTIAL).


```
FT TRANSMEM 215 235 7 (POTENTIAL).
FT DOMAIN 236 245 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 246 266 8 (POTENTIAL).
FT DOMAIN 267 279 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 280 299 9 (POTENTIAL).
FT DOMAIN 300 306 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 307 327 10 (POTENTIAL).
FT DOMAIN 328 342 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 343 363 11 (POTENTIAL).
FT DOMAIN 364 370 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 371 391 12 (POTENTIAL).
FT DOMAIN 392 419 EXTRACELLULAR (POTENTIAL).
SQ SEQUENCE 419 AA; 44243 MW; B5BF1D97DD7B9798 CRC64;

Query Match 8.8%; Score 79.5; DB 1; Length 419;
Best Local Similarity 23.2%; Pred. No. 6.5;
Matches 32; Conservative 21; Mismatches 38; Indels 47; Gaps 6;

QY 50 FLNLFGLFGIGSF-----AOGDILG-----GFLILGFDVAGVIGLILTGAVLDIKAL 96
Db 216 WLVTLCVAACMGSEFFVPEFIAPGLMGRQGVSQLGFSLL-FAIVAIAMVETAREMG---- 270
97 DKNAPRAAFKW-----TWGKGMLAGAVTMVTRLTETIIIPFTFANSYRKLNKSNLNIA 150
Db 271 -----RVIPKWSPSVLRMGMGCLAGAVLLAITEIW-----ALQSV 307
QY 151 FGGFEPSEFDINMGQASAL 168
Db 308 LGFIAPMWLVIGVATRAV 325

RESULT 11
ID DCTA_PSEAE STANDARD; PRT; 436 AA.
AC Q914F5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE C4-dicarboxylate transport protein.
GN DCTA OR PA1183.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=ATCC 15692 / PA01;
RA MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -!- FUNCTION: Responsible for the transport of dicarboxylates such as
succinate, fumarate, and malate from the periplasm across the
inner membrane (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(BY similarity).
CC -!- SIMILARITY: BELONGS TO THE SODIUM:DICARBOXYLATE SYMPORTER FAMILY
(SDF, TC 2.A.23).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
```

```
DR EMBL; AE004548; AAG04572.1; -.
DR InterPro; IPR001991; Na/dico_symp.
DR Pfam; PF00375; SDF; 1.
DR PRINTS; PR00173; EDTRNSPORT.
DR PROSITE; PS00713; NA_DICARBOXYL_SYMPT_1; FALSE_NEG..
DR PROSITE; PS00714; NA_DICARBOXYL_SYMPT_2; 1.
KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
Complete proteome.
FT TRANSMEM 9 28 POTENTIAL.
FT TRANSMEM 48 65 POTENTIAL.
FT TRANSMEM 78 100 POTENTIAL.
FT TRANSMEM 141 163 POTENTIAL.
FT TRANSMEM 189 211 POTENTIAL.
FT TRANSMEM 226 248 POTENTIAL.
FT TRANSMEM 290 312 POTENTIAL.
FT TRANSMEM 327 349 POTENTIAL.
FT TRANSMEM 354 376 POTENTIAL.
SQ SEQUENCE 436 AA; 45993 MW; DCA564DB31A5AC07 CRC64;

Query Match 8.7%; Score 78.5; DB 1; Length 436;
Best Local Similarity 37.1%; Pred. No. 8.2;
Matches 26; Conservative 7; Mismatches 24; Indels 13; Gaps 3;

QY 22 DSKSAFNLGAGEKLLAYETSKKDPVIFPLNLFGLFGIGSFAQGDILGFLILGFDVAV- 80
Db 114 DTSKIAAYAAAGE-----KQSTVDLMNVIPGTGVGAFANGDIL---QVLFFSVLF 161
QY 81 GIGLILTGAY 90
Db 162 GYALHRLGSY 171

RESULT 12
ID CYOB_ECOLI STANDARD; PRT; 663 AA.
AC P18401;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquinol oxidase polypeptide I (EC 1.10.3.-) (Cytochrome O subunit 1)
DE (Ubiquinol oxidase chain A).
DE (Cytochrome O subunit 1) (Cytochrome O ubiquinol oxidase subunit 1)
GN CYOB OR B0431 OR Z0534 OR ECS0485.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RC SEQUENCE FROM N.A.
RX MEDLINE=90293062; PubMed=2162835;
RA Chepur V., Lemieux L., Au D.C.T., Gennis R.B.;
RT "The sequence of the cyo operon indicates substantial structural
similarities between the cytochrome o ubiquinol oxidase of
Escherichia coli and the aa3-type family of cytochrome c oxidases.";
RL J. Biol. Chem. 265:11185-11192(1990).
RN [2]
RC SEQUENCE FROM N.A.
RX STRAIN=K12 / MGL1655;
RA MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RC SEQUENCE FROM N.A.
RA Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,
Duncan M., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
Lew H., Lin D., Namath A., Oefner P., Schramm S., Davis R.W.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
```


CCDA_BACSU STANDARD; PRT; 235 AA.

AC P45706;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Cytochrome c-type biogenesis protein ccda.

GN CCDA.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=97124194; PubMed=8969507;

RA Rose M., Entian K.D.;

RT "New genes in the 170 degrees region of the Bacillus subtilis genome encode DNA gyrase subunits, a thioedoxin, a xylanase and an amino acid transporter.";

RL Microbiology 142:3097-3101(1996).

[2]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=97221596; PubMed=9068642;

RA Schioett T., von Wachenfeldt C., Hederstedt L.;

RT "Identification and characterization of the ccda gene, required for cytochrome c synthesis in Bacillus subtilis.";

RL J. Bacteriol. 179:1962-1973(1997).

[3]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S., Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A., Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N., Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G., Guisepi G., Guy B.J., Haga K., Haeht J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K., Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H., Weitzrenger T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;

RT "The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.";

RL Nature 390:249-256(1997).

[4]

RP FUNCTION, AND SUBCELLULAR LOCATION.

RC STRAIN=168;

RX MEDLINE=2045543; PubMed=10781554;

RA Schioett T., Hederstedt L.;

RT "Efficient spore synthesis in Bacillus subtilis depends on the CcdA protein.";

RL J. Bacteriol. 182:2845-2854(2000).

CC -1- FUNCTION: REQUIRED FOR CYTOCHROME C SYNTHESIS AND STAGE V OF SPOULATION. MIGHT TRANSFER REDUCING EQUIVALENTS ACROSS THE

CC

CYTOPLASMIC MEMBRANE, PROMOTING EFFICIENT DISULFIDE BOND ISOMERIZATION OF PROTEINS LOCALIZED ON THE OUTER SURFACE OF THE MEMBRANE OR IN THE SPORE COAT.

-1- PATHWAY: CYTOCHROME C SYNTHESIS SYSTEM II.

-1- SUBCELLULAR LOCATION: Integral membrane protein.

-1- SIMILARITY: BELONGS TO THE DSBD SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; Z73234; CA97594.1; -

EMBL; X87845; CA61116.1; -

EMBL; 299113; CAB13677.1; -

Subtilist; EGI1240; ccda.

InterPro; IPR003834; Ctytoch_TM.

Pfam; PF02683; Dsbd; 1.

Cytochrome c-type biogenesis; Sporulation; Transmembrane; Complete proteome.

TRANSMEM 10 30 POTENTIAL.

TRANSMEM 57 77 POTENTIAL.

TRANSMEM 87 107 POTENTIAL.

TRANSMEM 133 153 POTENTIAL.

TRANSMEM 165 185 POTENTIAL.

TRANSMEM 204 224 POTENTIAL.

SEQUENCE 235 AA; 26007 MW; D94F3E62EDD2AC29 CRC64;

Query Match 8.6%; Score 78; DB 1; Length 235;

Best Local Similarity 23.5%; Pred. No. 5;

Matches 43; Conservative 25; Mismatches 67; Indels 48; Gaps 8;

QY 11 AFCVFSSQAQADSKSAFNLGAGEKLLAYETSKDPIVPFLN---LFLGFEGISFAQGD 67

|| : || : || : || : || : || : || : || : || : || : || : || : || : ||

Db 28 AFLSVITGVSMDDVKT-----EKLQKRSFLHTLC-FLLGFSVITIALGYGTSFGS 79

QY 68 I-----LGGFLILGFDVAGTGLTLTGAYLDIKALD-KNAPKA-----AF 105

|| : || : || : || : || : || : || : || : || : || : || : || : || : ||

Db 80 LFRDYHDAIRIGALLILFGITLGVPRPEAMKERRIHKPKSPGSLVGLIGMAFAA 139

QY 106 KWTGKGMLAGAVTMAVTR-----LTEIIIPFTFANSYNRKL-----KNLSLIA 150

|| : || : || : || : || : || : || : || : || : || : || : || : || : ||

Db 140 GWTCTGPILAIVITLACTNPGSAVPYMWLVGLFAVDFLLSFTITLKWIRKNQLFIM 199

QY 151 FGG 153

Db 200 KAG 202

RESULT 14

DCDA_RHILE

ID DCDA_RHILE STANDARD; PRT; 444 AA.

AC Q01857;

DT 01-APR-1993 (Rel. 25, Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE C4-dicarboxylate transport protein.

GN DCDA.

OS Rhizobium leguminosarum.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Rhizobium.

OX NCBI_TaxID=384;

[1]

RP SEQUENCE FROM N.A.

RX Ronson C.W.;

RL Submitted (DCC-1991) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Responsible for the transport of dicarboxylates such as succinate, fumarate, and malate from the periplasm across the inner membrane. This transport system plays an important role in the energy supply of rhizobium-legume symbionts (By similarity).

CC

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SODIUM:DICARBOXYLATE SYMPORTER FAMILY
CC (SDF, TC 2.A.23).
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z11529; CAA77618.1; .
DR PIR; S25701; S25701.
DR PIR; S27384; S27384.
DR InterPro; IPR001991; Na/dico_symp.
DR Pfam; PF00375; SDF; 1.
DR PROSITE; PS00713; NA_DICARBOXYL_SYMPT_1; 1.
DR PROSITE; PS00714; NA_DICARBOXYL_SYMPT_2; 1.
KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport.
FT TRANSMEM 18 40 POTENTIAL.
FT TRANSMEM 53 75 POTENTIAL.
FT TRANSMEM 90 112 POTENTIAL.
FT TRANSMEM 142 159 POTENTIAL.
FT TRANSMEM 163 180 POTENTIAL.
FT TRANSMEM 201 222 POTENTIAL.
FT TRANSMEM 232 254 POTENTIAL.
FT TRANSMEM 327 349 POTENTIAL.
FT TRANSMEM 364 386 POTENTIAL.
SQ SEQUENCE 444 AA; 46044 MW; AA2A5EA9B6780653 CRC64;

Query Match 8.6%; Score 77.5; DB 1; Length 444;
Best Local Similarity 29.1%; Pred No. 10;
Matches 44; Conservative 15; Mismatches 47; Indels 45; Gaps 9;
QY 38 AYETSKDPVFLNLFGLFGIGSFAQGDILGGLGFDVAV-GIGLILTG----- 88
Db 137 AHEQS-----IVGFLNIIPITTVGAFADGDIL---QVLFFSVLFGIALAMVGEKGEQVYN 189
QY 89 -----AYLDIKALDKNPKAAAF---KWTWCK-GM-----MLAGAVTNVAVTLRLTEI 129
Db 190 FLNLSLTPVFKVLVAILMKAAPIGAGMAFTTIGKYGVSIGANLMLIG--TFYITSLLEFV 247
QY 130 IIPFTFANSYN-----RLKNSLNIAFG 152
248 FIVLGAVARYNGFSIVALLRYIKELELLVIG 278

RESULT 15
SLAP_CAMPE
ID SLAP_CAMPE STANDARD; PRT; 933 AA.
AC P35827;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE S-layer protein (Surface array protein) (SAP).
GN SAP.
OS Campylobacter fetus.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=196;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.
RC STRAIN=84-32 / 23D;
RX MEDLINE=90354448; PubMed=2387868;
RA Blaser M.J., Gotschlich E.C.;
RT "Surface array protein of Campylobacter fetus. Cloning and gene
RT structure."
RL J. Biol. Chem. 265:14529-14535(1990).
RN [2]
RP ERRATUM.
RX MEDLINE=91035477; PubMed=2229082;

RA Blaser M.J., Gotschlich E.C.;
RL J. Biol. Chem. 265:19372-19372(1990).
CC -!- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
CC OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA. THIS PROTEIN IS
CC CRITICAL FOR VIRULENCE.
CC -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC S-LAYER WITH HEXAGONAL SYMMETRY.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J05577; AAA23032.1; .
KW Cell wall; S-layer.
SQ SEQUENCE 933 AA; 96757 MW; F88C729B4BA5B1E9 CRC64;
Query Match 8.5%; Score 76.5; DB 1; Length 933;
Best Local Similarity 24.6%; Pred. No. 26;
Matches 42; Conservative 21; Mismatches 73; Indels 35; Gaps 6;
QY 14 VFSSFAQADDSKSAFN---LGAGEKLLAYETSKDPIVPFLNLFGLFGIGSFAQGDIL 69
Db 615 ISDAVATADLSSSAFNKSVIITTKAAADTLTKDKQVIN---FTAADAGSVK----- 664
QY 70 GGFLILGFDVAVGIGLILTGAYLDIKALDKN-----APKAAFKWTWVGKMLAGAVTMA 122
Db 665 --LITVVKLNDVTALMIVVKIVLDAAAKDTNIALGTAADKAL-----VIDTGTETLN 714
QY 123 VTRLTEIIPFTFANSYNRLKNSLNIAFGGPEPSFDINMGASALGFELS 173
Db 715 IISLVKATSPETTANIVNAKLTDVTISIIDGMQ-----ITLGRAGTAGTDYS 761

Search completed: March 14, 2003, 14:24:56
Job time : 10.3333 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:22:51 ; Search time 24.3333 Seconds
(without alignments)
1507.249 Million cell updates/sec

Title: US-09-508-487-21
Perfect score: 904
Sequence: 1 MNKFLIVLLAFVCFSSFAQ.....DINMGQASALGFELSPFKSY 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 904 | 100.0 | 178 | 2 Q9X3U9 | Q9x3u9 borrelia af |
| 2 | 786.5 | 87.0 | 177 | 2 Q9X3V0 | Q9x3v0 borrelia ga |
| 3 | 769.5 | 85.1 | 179 | 16 O51065 | O51065 borrelia bu |
| 4 | 293.5 | 32.5 | 161 | 16 O50896 | O50896 borrelia bu |
| 5 | 285 | 31.5 | 170 | 16 O50885 | O50885 borrelia bu |
| 6 | 240 | 26.5 | 190 | 16 O50696 | O50696 borrelia bu |
| 7 | 237 | 26.2 | 161 | 2 Q9S011 | Q9s011 borrelia bu |
| 8 | 160 | 17.7 | 95 | 2 O31325 | O31325 borrelia bu |
| 9 | 95.5 | 10.6 | 574 | 16 O51143 | O51143 borrelia bu |
| 10 | 92 | 10.2 | 427 | 2 Q93RE6 | Q93re6 vibrio para |
| 11 | 90 | 10.0 | 325 | 16 Q8REW2 | Q8rew2 fusbacteri |
| 12 | 90 | 10.0 | 479 | 2 Q9FDM0 | Q9fdm0 zymomonas m |
| 13 | 89.5 | 9.9 | 469 | 16 Q97JE7 | Q97je7 clostridium |
| 14 | 88 | 9.7 | 312 | 16 Q9RTD0 | Q9rtd0 deinococcus |
| 15 | 87.5 | 9.7 | 616 | 16 Q8UAV6 | Q8uav6 agrobacteri |
| 16 | 85.5 | 9.5 | 482 | 16 Q9ZMT7 | Q9zmt7 helicobacte |

| | | | | | | |
|----|------|-----|------|----|--------|--------------------|
| 17 | 85.5 | 9.5 | 538 | 17 | Q9HHI2 | Q9hhI2 halobacteri |
| 18 | 85 | 9.4 | 363 | 16 | Q929N4 | Q929n4 listeria in |
| 19 | 85 | 9.4 | 364 | 16 | Q8Y5D0 | Q8y5d0 listeria mo |
| 20 | 84 | 9.3 | 1606 | 3 | Q9UW87 | Q9uw87 candida alb |
| 21 | 83.5 | 9.2 | 650 | 16 | Q97RE5 | Q97re5 streptococc |
| 22 | 83.5 | 9.2 | 709 | 16 | P73100 | P73100 synecocyst |
| 23 | 83 | 9.2 | 445 | 16 | Q9KBP3 | Q9kbp3 bacillus ha |
| 24 | 82.5 | 9.1 | 559 | 5 | Q24796 | Q24796 echinococcu |
| 25 | 82.5 | 9.1 | 559 | 5 | Q05768 | Q05768 echinococcu |
| 26 | 81.5 | 9.0 | 442 | 17 | Q976V7 | Q976v7 sulfolobus |
| 27 | 81 | 9.0 | 214 | 16 | O55742 | O55742 synecocyst |
| 28 | 81 | 9.0 | 305 | 16 | O25355 | O25355 helicobacte |
| 29 | 80.5 | 8.9 | 357 | 16 | O66804 | O66804 aquifex ao |
| 30 | 80.5 | 8.9 | 661 | 10 | Q9XGQ5 | Q9xgq5 oryza sativ |
| 31 | 80 | 8.8 | 183 | 16 | Q9XIG6 | Q9xig6 thermotoga |
| 32 | 80 | 8.8 | 336 | 16 | P73160 | P73160 synecocyst |
| 33 | 80 | 8.8 | 379 | 16 | Q92L18 | Q92l18 rhizobium m |
| 34 | 80 | 8.8 | 518 | 16 | Q8XZV3 | Q8xzv3 ralstonia s |
| 35 | 80 | 8.8 | 779 | 16 | Q92FB1 | Q92fb1 listeria in |
| 36 | 79.5 | 8.8 | 419 | 2 | Q93G9 | Q93g9 pseudomonas |
| 37 | 79.5 | 8.8 | 517 | 16 | Q8XZC8 | Q8xzc8 ralstonia s |
| 38 | 79.5 | 8.8 | 559 | 5 | O24788 | O24788 echinococcu |
| 39 | 79.5 | 8.8 | 904 | 12 | Q9IWA2 | Q9iwa2 wheat yello |
| 40 | 79 | 8.7 | 298 | 2 | Q9FB74 | Q9fb74 helicobacte |
| 41 | 79 | 8.7 | 399 | 2 | Q9ANV6 | Q9anv6 lactobacill |
| 42 | 79 | 8.7 | 419 | 2 | Q9KJY5 | Q9kjiy5 klebsiella |
| 43 | 79 | 8.7 | 429 | 16 | Q9KFP6 | Q9kfp6 bacillus ha |
| 44 | 79 | 8.7 | 433 | 17 | Q8U2X3 | Q8u2x3 pyrococcus |
| 45 | 79 | 8.7 | 539 | 10 | O80918 | O80918 arabidopsi |

ALIGNMENTS

RESULT 1

Q9X3U9 ID Q9X3U9 PRELIMINARY; PRT; 178 AA.
AC Q9X3U9;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Membrane protein P13.
GN P13.
OS Borrelia afzelii.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_taxid=29518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ACAI;
RX MEDLINE=21189251; PubMed=11292755;
RA Noppa L., Ostberg Y., Lavrinovich M., Bergstrom S.;
RT "P13, an integral membrane protein of Borrelia burgdorferi, is C-terminally processed and contains surface-exposed domains."
RL Infect. Immun. 69:3323-3334(2001).
DR EMBL: AF085740; AAD28361.1; -;
SQ SEQUENCE 178 AA; 19194 MW; C1C032EB55320C86 CRC64;

Query Match 100.0%; Score 904; DB 2; Length 178;
Best Local Similarity 100.0%; Pred. No. 5.2e-74;
Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNKFLIVLLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKDPVPPVFLNLFGLFGFI 60
|||||

Db 1 MNKFLIVLLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKDPVPPVFLNLFGLFGFI 60
|||||

Qy 61 GSFAQGDILGGFLLILGFDVAVGILITGAYLDIKALDKNAPKAFKWTGKGMMLAGAVT 120
|||||

Db 61 GSFAQGDILGGFLLILGFDVAVGILITGAYLDIKALDKNAPKAFKWTGKGMMLAGAVT 120
|||||

Qy 121 MAVFRLTEIIIPFTFANSYNRKLNLSIAFGGFPSPFDINMGQASALGFELSPFKSY 178
|||||

Db 121 MAVFRLTEIIIPFTFANSYNRKLNLSIAFGGFPSPFDINMGQASALGFELSPFKSY 178
|||||

```
RESULT 2
Q9X3V0
ID Q9X3V0 PRELIMINARY; PRT; 177 AA.
AC Q9X3V0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Membrane protein P13.
GN P13.
OS Borrelia garinii.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=29519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IP90;
RC MEDLINE=21189251; PubMed=11292755;
RA Noppa L., Ostberg Y., Lavrinovich M., Bergstrom S.;
RA "P13, an integral membrane protein of Borrelia burgdorferi, is C-terminally processed and contains surface-exposed domains.";
RL Infect. Immun. 69:3323-3334(2001).
RL EMBL: AF085741; AAD28362.1; -.
SQ SEQUENCE 177 AA; 19308 MW; 429928CC426E1111 CRC64;

Query Match 87.0%; Score 786.5; DB 2; Length 177;
Best Local Similarity 86.5%; Pred. No. 2.1e-63;
Matches 154; Conservative 7; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNKFLIVLLAFVCFSSFAQADDSKSAFNLAGAGEKLLAYETSKKDPVFPFLLNLFGLGFI 60
DB 1 MNKFLIFVLCFVCFSSFAQADDSKSTFNLAGAGEKLVYETNKKDSLVFPFLLNLFGLGFI 59
QY 61 GSFAQGDILGGFLILGFDVAGVIGLITGAYLDIKALDKNAPKAAFWTKGKGMMLAGAVT 120
DB 60 GSFAQGDILGGSLILGFDVAGVIGLITGAYLDIKDFDNNAKADKFWTKGKGMMLAGVVT 119
QY 121 MAVTRLTEIIPFTFANSYNRKLNLSNIAFGFEPFSDINMGQASALGFELSPKSY 178
DB 120 MAVTRLTEIVLPFTFANNYNRKLNLSNIALGGFEPFSDINMGQASALGFELSPKSY 177

RESULT 3
O51065
ID O51065 PRELIMINARY; PRT; 179 AA.
AC O51065;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein BB0034.
GN BB0034 OR P13.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RC MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RA "Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.";
RA Nature 390:580-586(1997).
RL EMBL: AF000790; AAC66226.1; -.
RN [1]
RP SEQUENCE 161 AA; 17887 MW; D93BDB326FE2DA30 CRC64;

Query Match 32.5%; Score 293.5; DB 16; Length 161;
Best Local Similarity 41.1%; Pred. No. 6.1e-19;
Matches 74; Conservative 24; Mismatches 61; Indels 21; Gaps 6;

QY 1 MNKFLIVLLAFVCFSSFAQADDSKSAFNLAGAGEKLLAYETSKKDPVFPFLLNLFGLGFI 60
DB 1 MKKIFTLILFGLTIQIFA-TKDTQNRKIEGI-ESEFNKYDKENKPNIGPPELNLFPFGI 58
QY 61 GSFAQGDILGGFLILGFDVAGVIGLITGAYLDIKALDKNAPKAAFWTKGKGMMLAG-AV 119
DB 59 GSFAQGDILGGSLVGLGFLILGAILGTGTLN-----HRETLTGTILGVGA 106
QY 120 MAVTRLTEIIPFTFANSYNRKLNLSNIAFGFEPFSDINMGQASALGFELSPKSY 178
DB 107 SWLTSTVSLIIPFTFANNHNENLKRUSAEIAGFEPNFDLGIN-----GFOLSPKSY 161
```

```
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE001117; AAC66426.1; -.
DR EMBL: AF085739; AAD28360.1; -.
DR TIGR: BB0034; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 179 AA; 19104 MW; FD6056E8E24E5D5D CRC64;

Query Match 85.1%; Score 769.5; DB 16; Length 179;
Best Local Similarity 84.9%; Pred. No. 7.3e-62;
Matches 152; Conservative 10; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNKFLIVLLAFVCFSSFAQADDSKSAFNLAGAGEKLLAYETSKKDPVFPFLLNLFGLGFI 59
DB 1 MNKLLIFVLATFCVCFSSFAQADDSKAGFCMSAGEKLVYETSKQDDVFPFLLNLFGLGFI 60
QY 60 GSFAQGDILGGFLILGFDVAGVIGLITGAYLDIKALDKNAPKAAFWTKGKGMMLAGAV 119
DB 61 GSFAQGDILGGSLILGFDVAGVIGLITGAYLDIKALDKGTTKKAFFQMTWKGVMYLAGV 120
QY 120 TMAVTRLTEIIPFTFANSYNRKLNLSNIAFGFEPFSDINMGQASALGFELSPKSY 178
DB 121 TMAVTRLTEIIPFTFANSYNRKLNLSNIALGGFEPFSDVAMQSSALGFELSPKSY 179
```

```
RESULT 4
O50896
ID O50896 PRELIMINARY; PRT; 161 AA.
AC O50896;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein BBA01.
GN BBA01.
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp54
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RC MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RA "Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.";
RA Nature 390:580-586(1997).
RL EMBL: AE000790; AAC66226.1; -.
RN [1]
RP SEQUENCE 161 AA; 17887 MW; D93BDB326FE2DA30 CRC64;

Query Match 32.5%; Score 293.5; DB 16; Length 161;
Best Local Similarity 41.1%; Pred. No. 6.1e-19;
Matches 74; Conservative 24; Mismatches 61; Indels 21; Gaps 6;

QY 1 MNKFLIVLLAFVCFSSFAQADDSKSAFNLAGAGEKLLAYETSKKDPVFPFLLNLFGLGFI 60
DB 1 MKKIFTLILFGLTIQIFA-TKDTQNRKIEGI-ESEFNKYDKENKPNIGPPELNLFPFGI 58
QY 61 GSFAQGDILGGFLILGFDVAGVIGLITGAYLDIKALDKNAPKAAFWTKGKGMMLAG-AV 119
DB 59 GSFAQGDILGGSLVGLGFLILGAILGTGTLN-----HRETLTGTILGVGA 106
QY 120 TMAVTRLTEIIPFTFANSYNRKLNLSNIAFGFEPFSDINMGQASALGFELSPKSY 178
DB 107 SWLTSTVSLIIPFTFANNHNENLKRUSAEIAGFEPNFDLGIN-----GFOLSPKSY 161
```



```
ID O31325 PRELIMINARY; PRT; 95 AA.
AC O31325;
DT 01-JAN-1998 (TremBLrel. 05, Created)
DT 01-JAN-1998 (TremBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE (clone 8) s3.
GN S3.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N40;
RX MEDLINE=96305348; PubMed=8688460;
RA Feng S., Das S., Barthold S.W., Fikrig E.;
RT "Characterization of two genes, p11 and p5, on the Borrelia
   burgdorferi 49-kilo base linear plasmid.";
RT Biochim. Biophys. Acta 1307:270-272(1996).
EMBL; L41151; AAB62281.1; -.
SEQUENCE 95 AA; 10739 MW; D12F6791F8920C3A CRC64;

Query Match 17.7%; Score 160; DB 2; Length 95;
Best Local Similarity 44.4%; Pred. No. 3.8e-07;
Matches 36; Conservative 12; Mismatches 31; Indels 2; Gaps 2;

QY 1 MNKFLIVLLAFRCVSSFAQADDSKSAFNLAGKLLAYETSKDPIVFPFLNLFGLFGI 60
DB 1 MKKIFLILFLGLTIEIFA-TDQTQRIEKG-I-ESFNKYDKKKNPIGPFLLNLFPLFGI 58

QY 61 GSFAGDILGGFLILGFDVANG 81
DB 59 GSFVQGDYIGGGVILGNLLG 79

RESULT 9
OS1143 PRELIMINARY; PRT; 574 AA.
AC OS1143;
DT 01-JUN-1998 (TremBLrel. 06, Created)
DT 01-JUN-1998 (TremBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE PTS system, maltose and glucose-specific IIBC component (MALX).
GN B00116.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Castlens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Utterback T., Wattley L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
   burgdorferi.";
RL Nature 390:580-586(1997).
DR EMBL; AE001124; AAC66516.1; -.
DR HSP; P05053; 1IBA.
DR TIGR; BB0116; -.
DR InterPro; IPR001996; PTS_EIIB.
DR InterPro; IPR003352; PTS_EIIC.
DR Pfam; PF00367; PTS_EIIB; 1.
DR Pfam; PF02378; PTS_EIIC; 1.
KW Complete proteome.
SQ SEQUENCE 574 AA; 63513 MW; 5CB04166E9DA8D67 CRC64;

Query Match 10.6%; Score 95.5; DB 16; Length 574;
Best Local Similarity 25.5%; Pred. No. 1.9;
```

```
Matches 55; Conservative 25; Mismatches 67; Indels 69; Gaps 12;

QY 6 IVVLLAFRCVSSFAQADDSKSAFNLAGKLLAYETSKDPIVFPFLNLFGLFGI 47
DB 188 IVIILPFCVFLAIFCLIWSSF---DDLIALGLFVRFEYFGSFLYGLNRLLLPLGLH 244
QY 48 ----VPPF-----LLNLF-----LGFIGGSPAQG--DILGGF-LILGFD 78
DB 245 SILSPFPEFTSLGGVEIVNGDVRLGNIFYAQLDPSLGKSSGFAKISSGFYLSIMEG 304
QY 79 AVGIGLILTGAYLDIKALDKNAKPAKWTGKGMMLAGAVTMVTRLEIIPPTFANS 138
DB 305 LFGAAL---GVYKGVIVHEDKNVAA-----LFGALTAFLTGITE---PLEFLFI 349
QY 139 YNRKLNKLSNIAFGGF-----EPSFDINMGQASALGF 170
DB 350 FTAPLLLYFVHAAYSGFALLANFFNVITIGNSPSTGF 385

RESULT 10
Q93RE6 PRELIMINARY; PRT; 427 AA.
AC Q93RE6;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Adhesin.
GN VAG.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RA Akeda Y., Honda T.;
RT "Vibrio parahaemolyticus adhesin gene.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047560; BAB59008.1; -.
DR InterPro; IPR002528; MatE.
DR Pfam; PF01554; UPF0013; 2.
DR TIGRFAMs; TIGR00797; matE; 1.
SQ SEQUENCE 427 AA; 45909 MW; 0B608B9DC70E3C71 CRC64;

Query Match 10.2%; Score 92; DB 2; Length 427;
Best Local Similarity 25.4%; Pred. No. 2.9;
Matches 50; Conservative 25; Mismatches 68; Indels 54; Gaps 10;

QY 5 LIVVLLAFRCVSSFAQ-----ADDSKSAFNLAGKLLAYETSKDPIV----- 48
DB 73 MVFAALAVCLFVFPQIMGLTDSQEVIELGSSYLVISSASMFACVAVGLRAMH 132
QY 49 -PFLNLFGLFGIGS--FAQGDILGGFLILGFDVANG-----ICLILTGAYLDIKALDKNA 100
DB 133 QGLSTFFSGIGILSNVFLNWLIFGH--LGPALGITGAATVVISGA-IEVGCL----- 185
QY 101 PKAAFKWTW-----GKGMMLAGAVTMVTRLEIIPPTFANSYNRKLNKLSNIAFG 152
DB 186 ----FGYLMKLLKHIITAFGWDIRASLVLDKITRFLSLSLPTTF-----NFLAWAGG 232
QY 153 GFEPSPFDINMGQASALG 169
DB 233 LF--AYHAIMGQAGVQG 247

RESULT 11
Q8REW2 PRELIMINARY; PRT; 325 AA.
AC Q8REW2;
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE CblB protein.
GN FN0975.
OS Fusobacterium nucleatum (subsp. nucleatum).
```



```
OC Bacteria; Fusobacteria; Fusobacterium.
OX NCBI_TaxID=76856;
RN SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RA Bhatnagar V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal R.,
RA Larsen N., D'Souza M., Walunas T., Pusck G., Haselkorn R.,
RA Fonstein M., Kyrplides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RL nucleatum strain ATCC 25586."
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL: AE010604; AAL95171.1;
KW Complete proteome.
SQ SEQUENCE 325 AA; 36409 MW; 9A53AE0C089DBAEE CRC64;

Query Match 10.0%; Score 90; DB 16; Length 325;
Best Local Similarity 23.2%; Pred. No. 3.2;
Matches 49; Conservative 29; Mismatches 85; Indels 48; Gaps 9;

4 FLIVLLAFVFFSFAQADSKSAF-----NLGAGEKLLAY----- 39
Db 76 YVIEFFLTATKSLADEGKKVYRILKSGDIEKAKKELSYLVSRDNTLSLDKIIMSV 135
Qy 40 -ETSKKDPVFPFLNLFLGFGISFAQGDILGGFLILGFDAGVIGLILTGAYLDIKALD- 97
Db 136 VETAEIVDGFISPVFAF-VGSFFYVELFG-----KVVSALPFAFTYKAINTLDS 187
Qy 98 ----KNAPKAFAKFTWKGMMLAGVMTVTRTEIIPF-TFANSYNRKLNKLSNIAFG 152
Db 188 MVGYKNEKYIDFGKVSARVDDIANFIPARLTGL--IFVPLSLILGYN--FKNSLKIFFR 243
Qy 153 GFEPSPDINMQ-----ASALGPFLSEPKSY 178
Db 244 DRKHSSPNSQSESAYAGALGIGFGKISY 274

RESULT 12
Q9FDMO PRELIMINARY; PRT; 479 AA.
AC Q9FDMO;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Metabolite transport protein.
OS Zymomonas mobilis.
OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;
NCBI_TaxID=542;
RN SEQUENCE FROM N.A.
RC STRAIN=ZM4;
RA Lee H.J., Kang H.S.;
RT "Sequence analysis of fosmid clone 42D10 of Zymomonas mobilis ZM4."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL: AF212041; AAG02149.1;
DR InterPro: IPR003663; CHO transporter.
DR InterPro: IPR003662; sub transporter.
DR Pfam: PF000083; sugar_tr; 1.
DR PRINTS: PR00171; SUGRTRNSPORT.
DR TIGRFAMS: TIGR00879; SP; 1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
DR Sugar transport; Transmembrane.
KW SEQUENCE 479 AA; 52095 MW; 614B052B61AD2A12 CRC64;

Query Match 10.0%; Score 90; DB 2; Length 479;
Best Local Similarity 22.5%; Pred. No. 4.9;
Matches 29; Conservative 24; Mismatches 52; Indels 24; Gaps 4;
```

```
Qy 7 VVLLAFVFFSFAQADSKSAFNLGAGEKLLAYETSKKDPVFPFLNLFLGFGISFAQG 66
Db 98 MIMIVTAIFGSVIAADAPTAFNLGAAARLVLGFAVGSSQIVPVY-----TAEIAPA 149
Qy 67 DILGGFLILGFDAGVIGLI---LFGAYLDIKALDKNAPKAFAKFTWKGMMLAGVMTAV 123
Db 150 DQGRMVTFYNISIGILGAAGIVGAFIQ-----EETW-RTMFSVAIPAIV 196
Qy 124 TRLTEIIP 132
Db 197 LFCMMMLP 205

RESULT 13
Q97JE7 PRELIMINARY; PRT; 469 AA.
AC Q97JE7;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Possible sugar-proton symporter.
GN CAC1339.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1488;
RN SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21353325; PubMed=11466286;
RA Neelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RL bacterium Clostridium acetobutylicum."
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL: AE007645; AAK79307.1;
DR InterPro: IPR003663; CHO transporter.
DR InterPro: IPR003662; sub transporter.
DR Pfam: PF00083; sugar_tr; 1.
DR PRINTS: PR00171; SUGRTRNSPORT.
DR TIGRFAMS: TIGR00879; SP; 1.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 469 AA; 50482 MW; CDE33E2CD2A7B0DA CRC64;

Query Match 9.9%; Score 89.5; DB 16; Length 469;
Best Local Similarity 25.1%; Pred. No. 5.3;
Matches 44; Conservative 26; Mismatches 66; Indels 39; Gaps 8;

Qy 5 LIVVLLAFVFFSFAQADSKSAFNLGAGEKLLAYETSKKDPVFPFLNLFL-----LGF 58
Db 11 LLFTVLISCA-----AGIGLLYGYDTAVISGAIGFLKKLYNLSPAMQGF 55
Qy 59 GIGSFAQGDILG---GFLILGFDAGVIG--LIITGAYLDIKALDKNAPKAFAKFTWCK- 111
Db 56 VISSIMGGVILGVGFGSFL---GDAIGRRKVLMAALFAISAVISSISTSAFMLIFARI 112
Qy 112 --GMMLAGAVTMVTRTEIIPFTFANSYNRKLNKLSNLI-AFGGFEPSPDINMG 163
Db 113 VGGIGIGMASALSVTYITECAPP-----SIRGLSSLYQLFTILGISITFFPNLG 162

RESULT 14
Q9RTD0 PRELIMINARY; PRT; 312 AA.
AC Q9RTD0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Phospho-N-acetylmuramoyl-pentapeptide- transferase.
```

GN DR1835.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Mofat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarov K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
radiodurans RI.";
RT Science 286:1571-1577(1999).
EMBL; AE002024; AAF11390.1; -
TIGR; DR1835; -
DR InterPro: IPR00715; Glycos transf. 4.
DR InterPro: IPR003524; PNAcPP transf.
DR Pfam: PF00953; Glycos transf. 4; 1.
DR PROSITE: PS01347; MRAY_1; UNKNOWN_1.
DR PROSITE: PS01348; MRAY_2; UNKNOWN_1.
KW Transferrase; Complete proteome.
SQ SEQUENCE 312 AA; 33450 MW; 3A284B58E9395B2F CRC64;

Query Match 9.7%; Score 88; DB 16; Length 312;
Best Local Similarity 22.7%; Pred. No. 4.6;
Matches 40; Conservative 26; Mismatches 54; Indels 56; Gaps 7;

QY 4 FLIVLLAF---CVSSFAQADDSKSAFNL-----GAGEKLL 37
Db 51 FVLALLVFLGAFGGICQANLSREVILLALGMGVGGIDDFLKIRSRFGGKELL 110
QY 38 AYETSKDPVFPFLNLFGLFGIGSFAQGDILGGFLILG----FDVAGIGLILTGAYLDI 93
Db 111 ARE---KFPQLVALLFAGFA-APLASHQLPFGMSIGGYPIFDMLFAFVMGVS--- 163
QY 94 KALDKNAPKAAPKWTWCKGMVLG-----AVTMAVTRLTETIIIPFTFANSY 139
Db 164 -----NAPNFTDGLDGLAGVMIVLLPLVAVSPIALMLVAVLLGLFWFNAH 210

RESULT 15
Q8UAV6 PRELIMINARY; PRT; 616 AA.
AC Q8UAV6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ABC transporter, membrane spanning protein.
GN ATU3261 OR AGR_L_3113.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens

RT C58.";
RL Science 294:2317-2323(2001).
RN [2]
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmlel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursun J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AE009255; AAL44077.1; -
DR EMBL; AE008358; AAK90129.1; -
KW Complete proteome.
SQ SEQUENCE 616 AA; 63402 MW; A6005D4347ABE060 CRC64;

Query Match 9.7%; Score 87.5; DB 16; Length 616;
Best Local Similarity 22.0%; Pred. No. 11;
Matches 42; Conservative 35; Mismatches 71; Indels 43; Gaps 9;

QY 9 LIAFCVFSFAQADDSKSAFNLG-----AGEKLLAYETSKKDPV--- 49
Db 301 LLTSVLFSGAGSSVANAAFGASTFQPELVKHGYPQAQAGATIAA--TSVLDNVIPPSIA 358
QY 50 -FLNLFLGFGIGCSFAQGDILGGFLILGFDVAGIGLILTGAYLDIKALDKNAPKA--AFK 106
Db 359 FLILATATNLSVGS---LVGGFFAGGLMAVCLGVAI---HLSVRSD-TLPRATGAER 410
QY 107 W-----TWGKMMLAGAVTMAVTRLTETIIIPFTFANSYNRKLKNSLNIAFGGPEPSF 158
Db 411 WRSATAAIPAFGLGVIVVVGIRIGIVTTTEAA---ALAAALYLLGLGFGYRLGVRIFATF 467
QY 159 DINMGQASALG 169
Db 468 ROSAGEAAIG 478

Search completed: March 14, 2003, 14:27:04
JOB time : 26.3333 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:23:06 ; Search time 11 Seconds
(without alignments)
476.116 Million cell updates/sec

Title: US-09-508-487-21
Perfect score: 904
Sequence: 1 MNKFLIVLLAFVCSFSAQ.....DINMQASALGFELSFKRSY 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/FCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------------|-------------------|
| 1 | 81.5 | 9.0 | 858 | US-09-255-829-22 | Sequence 22, Appl |
| 2 | 81.5 | 9.0 | 858 | US-09-255-829-29 | Sequence 29, Appl |
| 3 | 81.5 | 9.0 | 1169 | US-09-255-829-20 | Sequence 20, Appl |
| 4 | 72 | 8.0 | 344 | US-09-134-001C-4175 | Sequence 4175, Ap |
| 5 | 72 | 8.0 | 451 | US-08-679-635A-2 | Sequence 2, Appl |
| 6 | 72 | 8.0 | 451 | US-09-419-163-2 | Sequence 2, Appl |
| 7 | 70.5 | 7.8 | 3491 | US-07-642-734C-2 | Sequence 2, Appl |
| 8 | 70.5 | 7.8 | 3491 | US-08-439-009A-2 | Sequence 2, Appl |
| 9 | 69.5 | 7.7 | 376 | US-08-608-241-2 | Sequence 2, Appl |
| 10 | 69.5 | 7.7 | 376 | US-08-922-182-2 | Sequence 2, Appl |
| 11 | 69.5 | 7.7 | 376 | US-08-919-953-2 | Sequence 2, Appl |
| 12 | 69.5 | 7.7 | 376 | US-09-192-983-2 | Sequence 2, Appl |
| 13 | 69.5 | 7.7 | 4928 | US-09-036-987A-5 | Sequence 5, Appl |
| 14 | 69.5 | 7.7 | 4928 | US-09-370-700-5 | Sequence 5, Appl |
| 15 | 69 | 7.6 | 931 | US-08-624-655A-2 | Sequence 2, Appl |
| 16 | 69 | 7.6 | 2285 | US-09-308-375-2 | Sequence 2, Appl |
| 17 | 68 | 7.5 | 280 | US-08-855-140-1 | Sequence 1, Appl |
| 18 | 67.5 | 7.5 | 646 | US-09-232-200-25 | Sequence 25, Appl |
| 19 | 67.5 | 7.5 | 646 | US-09-232-200-32 | Sequence 32, Appl |
| 20 | 67.5 | 7.5 | 646 | US-09-232-200-38 | Sequence 38, Appl |
| 21 | 67.5 | 7.5 | 646 | US-09-232-200-43 | Sequence 43, Appl |
| 22 | 67.5 | 7.5 | 646 | US-09-232-200-47 | Sequence 47, Appl |
| 23 | 67.5 | 7.5 | 646 | US-09-232-200-59 | Sequence 59, Appl |
| 24 | 67.5 | 7.5 | 646 | US-09-232-197-25 | Sequence 25, Appl |
| 25 | 67.5 | 7.5 | 646 | US-09-232-197-32 | Sequence 32, Appl |
| 26 | 67.5 | 7.5 | 646 | US-09-232-197-38 | Sequence 38, Appl |
| 27 | 67.5 | 7.5 | 646 | US-09-232-197-43 | Sequence 43, Appl |

| | | | | | | |
|----|------|-----|------|---|---------------------|-------------------|
| 28 | 67.5 | 7.5 | 646 | 4 | US-09-232-197-47 | Sequence 47, Appl |
| 29 | 67.5 | 7.5 | 646 | 4 | US-09-232-197-59 | Sequence 59, Appl |
| 30 | 67.5 | 7.5 | 646 | 4 | US-09-232-201-25 | Sequence 25, Appl |
| 31 | 67.5 | 7.5 | 646 | 4 | US-09-232-201-32 | Sequence 32, Appl |
| 32 | 67.5 | 7.5 | 646 | 4 | US-09-232-201-38 | Sequence 38, Appl |
| 33 | 67.5 | 7.5 | 646 | 4 | US-09-232-201-43 | Sequence 43, Appl |
| 34 | 67.5 | 7.5 | 646 | 4 | US-09-232-201-47 | Sequence 47, Appl |
| 35 | 67.5 | 7.5 | 646 | 4 | US-09-232-201-59 | Sequence 59, Appl |
| 36 | 67.5 | 7.5 | 934 | 1 | US-08-215-805A-80 | Sequence 80, Appl |
| 37 | 67.5 | 7.5 | 1584 | 4 | US-09-251-645-6 | Sequence 6, Appl |
| 38 | 67 | 7.4 | 226 | 4 | US-09-134-001C-5590 | Sequence 5590, Ap |
| 39 | 66.5 | 7.4 | 325 | 4 | US-09-134-001C-3513 | Sequence 3513, Ap |
| 40 | 66.5 | 7.4 | 646 | 4 | US-09-232-200-33 | Sequence 33, Appl |
| 41 | 66.5 | 7.4 | 646 | 4 | US-09-232-200-65 | Sequence 65, Appl |
| 42 | 66.5 | 7.4 | 646 | 4 | US-09-232-200-92 | Sequence 92, Appl |
| 43 | 66.5 | 7.4 | 646 | 4 | US-09-232-197-33 | Sequence 33, Appl |
| 44 | 66.5 | 7.4 | 646 | 4 | US-09-232-197-65 | Sequence 65, Appl |
| 45 | 66.5 | 7.4 | 646 | 4 | US-09-232-197-92 | Sequence 92, Ap |

ALIGNMENTS

RESULT 1
US-09-255-829-22
; Sequence 22, Application US/09255829
; Patent No. 6461617
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN, & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/255.829
; FILING DATE: 23-FEB-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB97/02273
; FILING DATE: 22-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/782.893
; FILING DATE: 27-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32.893
; REFERENCE/DOCKET NUMBER: 1581.0130002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 858 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-255-829-22

Query Match 9.0%; Score 81.5; DB 4; Length 858;
Best Local Similarity 23.8%; Pred. NO. 0.64;
Matches 36; Conservative 24; Mismatches 56; Indels 35; Gaps 6;


```

; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus aureus
; STRAIN: RUSA 315
US-08-679-635A-2

Query Match      8.0%; Score 72; DB 2; Length 451;
Best Local Similarity 21.3%; Pred. No. 3.5;
Matches 37; Conservative 28; Mismatches 71; Indels 38; Gaps 6;

QY 27 AFNLGA-GEKLLAYETSKDPVIFPFLNLFILGFCI---GSPAOGDILGGFLLILGFDVAVGI 82
    ||| ||| :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 23 AFKLGRTGVYLAHNKKEKHP-----RVLVGRDTRVSGEMLESALLIAGLISIGAEVMRL 76

QY 83 GLIIT-----GAYLDIKALDKKNAPKA--AFKWTWKGKMMLAGAVTMVTRLFEII 130
    ||: ||| :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 77 GIITPGVAYITRDMAELGYMISASHNPVADNGIKFSGDGFKLSDEQENEIEALLDQE 136

QY 131 IP-----FTFANSYNNKLNLSINIAFGFPEPSDINWGOASAL 168
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 137 NPFLPRPVGNDIVHYSDFEQAQKLYLSKLTVDVNFEGKLIALDANGANGSTSSL 190

RESULT 6
US-09-419-163-2
; Sequence 2, Application US/09419163
; Patent No. 6391614
; GENERAL INFORMATION:
; APPLICANT: Tomasz, Alexander
; APPLICANT: Delencastre, Herminia
; TITLE OF INVENTION: AUXILIARY GENES AND PROTEINS OF
; TITLE OF INVENTION: METHICILLIN RESISTANT BACTERIA AND ANTAGONISTS THEREOF
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patencin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/419,163
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/679,635
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-141
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus aureus
; STRAIN: RUSA 315
US-09-419-163-2

```

```

Db   1805 RHGVIEMP---PETACRALQNALDRA 1827
      :| | :| | :| | :| | :| | :| |
      :| | :| | :| | :| | :| | :| |

RESULT 8
US-08-439-009A-2
; Sequence 2, Application US/08439009A
; Patent No. 6004787
; GENERAL INFORMATION:
; APPLICANT: Donadio, S
; APPLICANT: Katz, L
; APPLICANT: Mcalpine, J B
; TITLE OF INVENTION: Method of Directing Biosynthesis of
; TITLE OF INVENTION: Specific Polyketides
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steven F. Weinstock
; STREET: Abbott Laboratories D377/AP6D-2 One Abbott
; STREET: Park Rd
; CITY: Abbott Park
; STATE: IL
; COUNTRY: US
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/439,009A
; FILING DATE: 11-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Casuto, Dianne
; REGISTRATION NUMBER: 40,943
; REFERENCE/DOCKET NUMBER: 4952.US.DI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-938-3137
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3491 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-439-009A-2

Query Match          7.8%; Score 70.5; DB 3; Length 3491;
Best Local Similarity 24.7%; Pred. No. 89;
Matches 36; Conservative 11; Mismatches 38; Indels 61; Gaps

QY    9 LLAFVFSFSAQADDSKSFANLGACEKLLEYTSKKDPIVPFLNLTFGFGIGSFAQGDI 68
      |||||
Db    1739 LTAFVLFFSFASA-----FCAPGLGGVAPGN- 1764
      |||||

QY    69 LGFLLIGFDVAVGIGLILTCAYLDIKALDKNA---PKAAFKW-TWGKGMMLAGAVTMAVT 124
      |||||
Db    1765 -----AYDLGLAQQRSDGLPATAVAWGTWAGSGMAEGAVADRFR 1804
      |||||

QY    125 RLTEIIIPFTFANSYNRKLKNSLNIA 150
      |||||
Db    1805 RHGVIEMP---PETACRALQNALDRA 1827
      :| | :| | :| | :| | :| | :| |
      :| | :| | :| | :| | :| | :| |

RESULT 9
US-08-608-241-2
; Sequence 2, Application US/08608241
; Patent No. 5747328
; GENERAL INFORMATION:
; APPLICANT: Donohue, Timothy J
; APPLICANT: Barber, Robert D
; APPLICANT: Witthuhn, Vernon
; TITLE OF INVENTION: MICROBIAL SYSTEM FOR FORMALDEHYDE
; TITLE OF INVENTION: SENSING AND REMEDIATION

```

```

;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/922,182
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/608,241
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27,386
; REFERENCE/DOCKET NUMBER: 960296.93511
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-922-182-2

Query Match 7.7%; Score 69.5; DB 2; Length 376;
Best Local Similarity 23.0%; Pred. No. 5.3;
Matches 43; Conservative 26; Mismatches 71; Indels 47; Gaps

QY 23 DKSASFNLGAGEKLLAY-----ETSKKDPVFPFLNMLFLGFGIG----- 61
DB 120 DGSFRESMLDGTPIILHYMGCSFTSVTLPEITAVAKVRPDAPFDKICYIGCGVYTTGIGAV 179
QY 62 -SPAQGDILGFLILGFDAVGI---GLILTGAYLDIKALDKNAPKAAFKWTMGK----- 111
DB 180 INTAKVEIGAKAVVFGIGGLNVIOGLKAGADMIIGVDLNNAK-----EWGERFGMT 234
QY 112 -----GMMLAGAVTMAVTRLTEI-IPETFANSYNRK-LKNSLNIAFGGFPSFDI 160
DB 235 HFVNPSEIDGDVVVAHLVNNTKTPFDQIGGADYTFDCTGNVKVMRQALEACHRGWGQSIVI 294
QY 161 NMCQASA 167
DB 295 GVAPAGA 301

RESULT 11
US-08-919-953-2
; Sequence 2, Application US/08919953
; Patent No. 5837481
; GENERAL INFORMATION:
; APPLICANT: Donohue, Timothy J
; APPLICANT: Barber, Robert D
; APPLICANT: Witthuhn, Vernon
; TITLE OF INVENTION: MICROBIAL SYSTEM FOR FORMALDEHYDE
; TITLE OF INVENTION: SENSING AND REMEDIATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/919,953
; FILING DATE:
; CLASSIFICATION: 435

```

```
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/608,241
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Seay, Nicholas J
;; REGISTRATION NUMBER: 27,386
;; REFERENCE/DOCKET NUMBER: 960296.93511
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 608-251-5000
;; TELEFAX: 608-251-9166
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 376 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-919-953-2

Query Match 7.7%; Score 69.5; DB 2; Length 376;
Best Local Similarity 23.0%; Pred. No. 5.3;
Matches 43; Conservative 26; Mismatches 71; Indels 47; Gaps 7;

QY 23 DSKSAPNLAGKLLAY-----ETSKKDPVPLNLFGLFGIG----- 61
DB 120 DTSRFSMLDGPILHYMGCSFNSVTVLPETAVAKVRPDAPFDKICYIGCGVTTGIGAV 179
QY 62 -SFAQGDILGGLILGFDVAGI-----GLILTGAYLIDIKALDNAPKAAFKWTWGK----- 111
DB 180 INTAKVEIGAKAVVFLGIGLNVIOQLKLAGADMIIGVDLNNNAK-----EMGERFGMT 234
QY 112 -----GMLAGAVTMAVTRLEI-IIPTEFANSYNRK-LKNSLNTAFGGFEPFSDI 160
DB 235 HFVNPSEIDGDVVVAHLVNNMTKTPFDQIGGADYTFDCTGNVKNVQRALEACHRGWGOSIVI 294
QY 161 NMQQASA 167
DB 295 GVAPAGA 301

RESULT 12
US-09-192-983-2
; Sequence 2, Application US/09192983A
; Patent No. 6242244
; GENERAL INFORMATION:
; APPLICANT: Donohue, Timothy
; APPLICANT: Barber, Robert
; APPLICANT: Witthuhn, Vernon
; TITLE OF INVENTION: Microbial System for Formaldehyde Sensing and
; TITLE OF INVENTION: Remediation
; FILE REFERENCE: 960296.95505
; CURRENT APPLICATION NUMBER: US/09/192,983A
; CURRENT FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/919,953
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 08/608,241
; EARLIER FILING DATE: 1996-02-28
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Rhodobacter sphaeroides
US-09-192-983-2

Query Match 7.7%; Score 69.5; DB 4; Length 376;
Best Local Similarity 23.0%; Pred. No. 5.3;
Matches 43; Conservative 26; Mismatches 71; Indels 47; Gaps 7;

QY 23 DSKSAPNLAGKLLAY-----ETSKKDPVPLNLFGLFGIG----- 61
DB 120 DTSRFSMLDGPILHYMGCSFNSVTVLPETAVAKVRPDAPFDKICYIGCGVTTGIGAV 179
QY 62 -SFAQGDILGGLILGFDVAGI-----GLILTGAYLIDIKALDNAPKAAFKWTWGK----- 111
```

```
DB 180 INTAKVEIGAKAVVFLGIGLNVIOQLKLAGADMIIGVDLNNNAK-----EMGERFGMT 234
QY 112 -----GMLAGAVTMAVTRLEI-IIPTEFANSYNRK-LKNSLNTAFGGFEPFSDI 160
DB 235 HFVNPSEIDGDVVVAHLVNNMTKTPFDQIGGADYTFDCTGNVKNVQRALEACHRGWGOSIVI 294
QY 161 NMQQASA 167
DB 295 GVAPAGA 301

RESULT 13
US-09-036-987A-5
; Sequence 5, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow Agrosciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,987A
; FILING DATE: 09-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)337-4816
; TELEFAX: (317)337-4847
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4928 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-036-987A-5

Query Match 7.7%; Score 69.5; DB 4; Length 4928;
Best Local Similarity 25.2%; Pred. No. 1.9e+02;
Matches 39; Conservative 25; Mismatches 48; Indels 43; Gaps 8;

QY 46 PIVPFLNLFGLFGIGS-----FAQGDILGGLILGFDFA-----VGIGLILTGAYLIDIKAL 96
DB 2231 PVPVPMVS-----GKTEALSQAQDALMSYLSNRVDASPRDIGYSLAVTRPALDHRAY 2283
QY 97 DKNAPKAAFKWTWGMMLAGAVTMAVTR-LFEII-----IPPTFANSYNR--KLKN 145
DB 2284 VLGDRAA-----LLPGLKALAVSNDAAEVITGTAAAGPVGVFSGQGWPMGS 2334
QY 146 SLNIAFGFEPF-----DINMGQASALGFELS 173
```


Db 2335 GLHSAPFVFADAFDEACCELDHGLQMARLRDVLS 2369

RESULT 14

US-09-370-700-5
; Sequence 5, Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIV1
; CURRENT APPLICATION NUMBER: US/09/370,700
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/36987
; EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 4928
; TYPE: PRT
; ORGANISM: Saccharopolyspora spinosa
US-09-370-700-5

Query Match 7.7%; Score 69.5; DB 4; Length 4928;
Best Local Similarity 25.2%; Pred. No. 1.9e+02;
Matches 39; Conservative 25; Mismatches 48; Indels 43; Gaps 8;

Qy 46 PIVPELNLNLFGLFGIGS-----FAQDILGGFLILGFDA---VIGLILTGAYLDIKAL 96
Db 2331 PVVFWMVVS-----GKTPEALSAQADALMSYLSNRVDASPRDIGYSLAVTRPALDHRV 2283
Qy 97 DKNAPKAAFKWTGKGMLAGAVTMAVTR-LTEII-----IPFTFANSYNR--KLKN 145
Db 2284 VLGDRAA-----LLPLKALAVSNDAAEVITGTRAGPVGVFVSGGGQWPGMG 2334
Qy 146 SLNTAFGGFEPF-----DINMQASALGFELS 173
Db 2335 GLHSAPFVFADAFDEACCELDHGLQMARLRDVLS 2369

RESULT 15

US-08-624-655A-2
; Sequence 2, Application US/08624655A
; Patent No. 6323005
; GENERAL INFORMATION:
; APPLICANT: DABAN, MONTERRAT
; APPLICANT: MEDRANO, ANDRES
; APPLICANT: ESPUNA, ENRIC
; APPLICANT: QUEROL, ENRIQUE
; TITLE OF INVENTION: TRANSFERRIN-BINDING PROTEIN 1 (Tbpl) GENE OF
; TITLE OF INVENTION: Actinobacillus pleuropneumoniae, ITS USE TO PREPARE
; TITLE OF INVENTION: PRODUCTS FOR THE UTILIZATION IN VACCINES FOR
; TITLE OF INVENTION: PLEUROPNEUMONIA AND AS DIAGNOSTIC REAGENTS
; FILE REFERENCE: P00740
; CURRENT APPLICATION NUMBER: US/08/624,655A
; CURRENT FILING DATE: 1996-03-22
; PRIOR APPLICATION NUMBER: 95 00592
; PRIOR FILING DATE: 1995-03-24
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 931
; TYPE: PRT
; ORGANISM: Actinobacillus pleuropneumoniae
US-08-624-655A-2

Query Match 7.6%; Score 69; DB 4; Length 931;

Best Local Similarity 21.2%; Pred. No. 21;
Matches 44; Conservative 27; Mismatches 69; Indels 68; Gaps 9;
Qy 9 LLAFCVFSSFAQADDSSKSAFNL-----GAGEKLLAYETSKKOPIV 48
Db 7 LISLALLSLFAVQSYAQAVOLNDVYVTGTTKKKAHKKENEVTGLGKVVKTPDTLSKEQV- 65
Qy 49 PFLNLNLFGLF-----GIGSFAQG-DILGGFLILGFDAVIGLILTG----- 88
Db 66 -----LGIRDLTRYDPGISVVEQGRGATTGYSIRGVDRNRVGLALDGLPQIQSYVSQY 118
Qy 89 -----AYLDIKALDNAPKAAFKWTGKGMMLAGAVTMAVTRLTETIIP----- 132
Db 119 SRSSSGAINEIEYENLRSI--QISKGASSSEFGSG-SLGGSVQFRTKEVSDIIRPGQSWG 175
Qy 133 FTFANSYNRKLKNSLN-IAFCGFEPSPD 159
Db 176 LDTKSAYSSKNQOOWLNSLAFAGTHNGFD 203
Search completed: March 14, 2003, 14:27:47
Job time : 16 secs

THIS PAGE BLANK (USPTO)